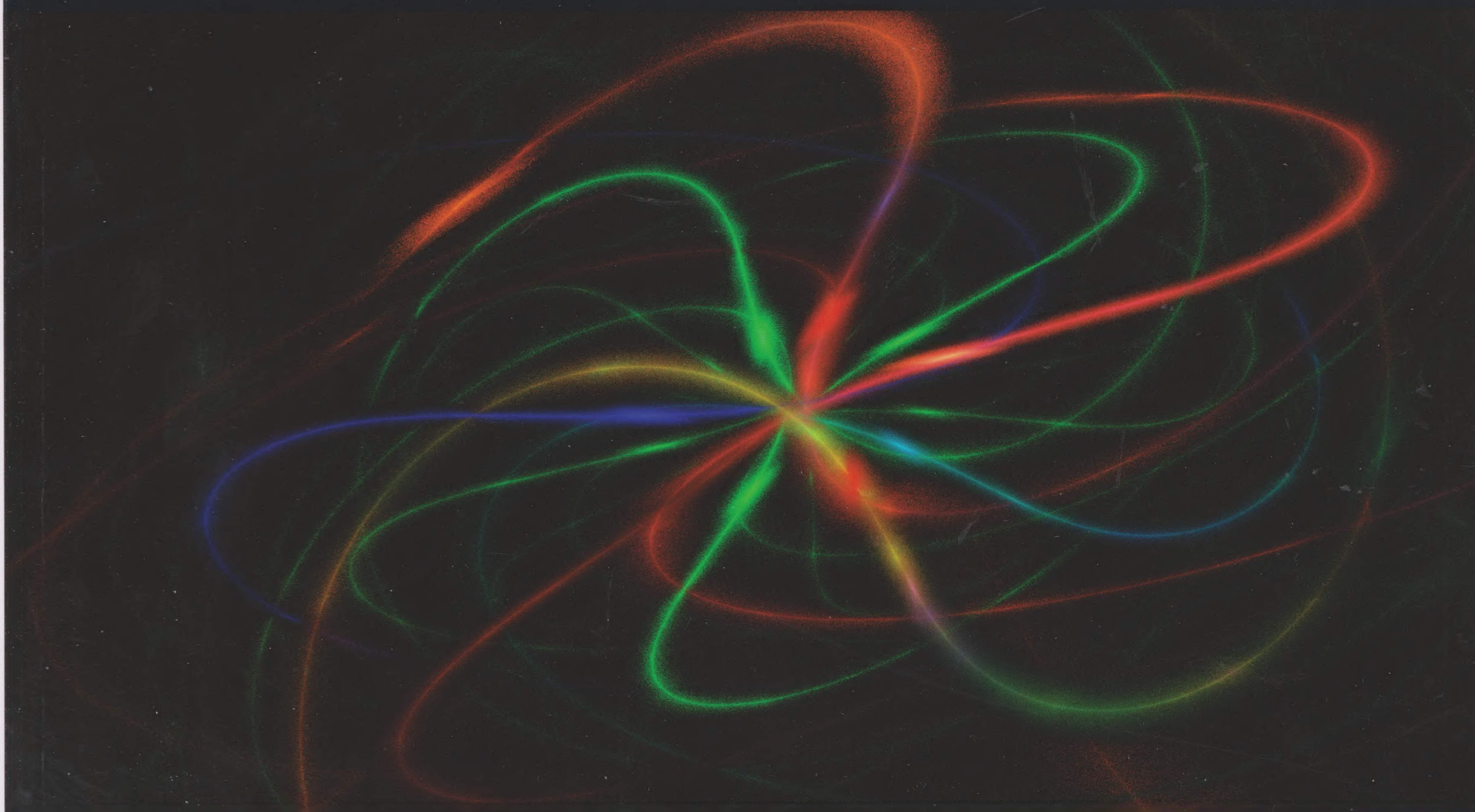
The Open University

MST125

Essential mathematics 2

Handbook



MST125

Essential mathematics 2

Handbook

This publication forms part of an Open University module. Details of this and other Open University modules can be obtained from the Student Registration and Enquiry Service, The Open University, PO Box 197, Milton Keynes MK7 6BJ, United Kingdom (tel. +44 (0)845 300 6090; email general-enquiries@open.ac.uk).

Alternatively, you may visit the Open University website at www.open.ac.uk where you can learn more about the wide range of modules and packs offered at all levels by The Open University.

To purchase a selection of Open University materials visit www.ouw.co.uk, or contact Open University Worldwide, Walton Hall, Milton Keynes MK7 6AA, United Kingdom for a brochure (tel. +44 (0)1908 858779; fax +44 (0)1908 858787; email ouw-customer-services@open.ac.uk).

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 2015. Second edition 2016.

Edited, designed and typeset by The Open University, using the Open University TEX System.

Printed in the United Kingdom by Latimer Trend and Company Ltd, Plymouth.

SUP 04700 7

2.1

Contents

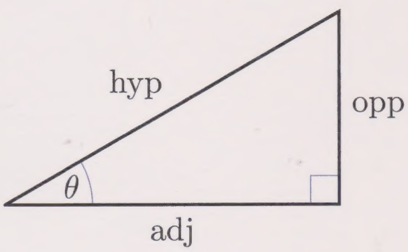
Quick reference material

Trigonometric ratios

Degrees and radians

$360^\circ = 2\pi$ radians $\qquad 1^\circ = \dfrac{\pi}{180}$ radians $\qquad 1$ radian $= \left(\dfrac{180}{\pi}\right)^\circ$

Sine, cosine and tangent

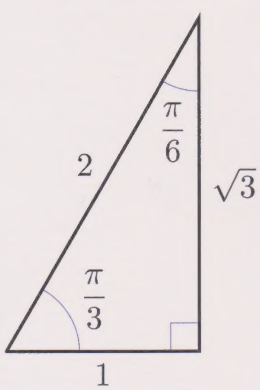


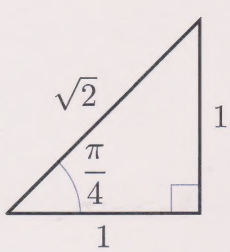
For an acute angle θ,

$$\sin\theta = \frac{\text{opp}}{\text{hyp}} \qquad \cos\theta = \frac{\text{adj}}{\text{hyp}} \qquad \tan\theta = \frac{\text{opp}}{\text{adj}}.$$

Mnemonic: SOH CAH TOA.

Special angles





θ in radians	θ in degrees	$\sin\theta$	$\cos\theta$	$\tan\theta$
0	0°	0	1	0
$\frac{\pi}{6}$	30°	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$
$\frac{\pi}{4}$	45°	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	1
$\frac{\pi}{3}$	60°	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$
$\frac{\pi}{2}$	90°	1	0	–
π	180°	0	-1	0

Note that $\dfrac{1}{\sqrt{2}} = \dfrac{\sqrt{2}}{2}$, and that $\tan\left(\dfrac{\pi}{2}\right)$ is undefined.

Trigonometric identities

tan, cosec, sec, cot in terms of sin and cos

$$\tan\theta = \frac{\sin\theta}{\cos\theta} \qquad \operatorname{cosec}\theta = \frac{1}{\sin\theta} \qquad \sec\theta = \frac{1}{\cos\theta} \qquad \cot\theta = \frac{\cos\theta}{\sin\theta}$$

Pythagorean identities

$$\sin^2\theta + \cos^2\theta = 1 \qquad \tan^2\theta + 1 = \sec^2\theta \qquad 1 + \cot^2\theta = \operatorname{cosec}^2\theta$$

Symmetry identities

$$\sin(-\theta) = -\sin\theta \qquad \sin(\theta + 2\pi) = \sin\theta \qquad \sin\left(\frac{\pi}{2} - \theta\right) = \cos\theta$$

$$\cos(-\theta) = \cos\theta \qquad \cos(\theta + 2\pi) = \cos\theta$$

$$\tan(-\theta) = -\tan\theta \qquad \tan(\theta + \pi) = \tan\theta \qquad \cos\left(\frac{\pi}{2} - \theta\right) = \sin\theta$$

Angle sum and angle difference identities

$$\sin(A+B) = \sin A\cos B + \cos A\sin B$$

$$\sin(A-B) = \sin A\cos B - \cos A\sin B$$

$$\tan(A+B) = \frac{\tan A + \tan B}{1 - \tan A\tan B}$$

$$\tan(A-B) = \frac{\tan A - \tan B}{1 + \tan A\tan B}$$

$$\cos(A+B) = \cos A\cos B - \sin A\sin B$$

$$\cos(A-B) = \cos A\cos B + \sin A\sin B$$

Double-angle and half-angle identities

$$\sin(2\theta) = 2\sin\theta\cos\theta$$

$$\begin{aligned}\cos(2\theta) &= \cos^2\theta - \sin^2\theta \\ &= 1 - 2\sin^2\theta \qquad \text{so} \quad \sin^2\theta = \tfrac{1}{2}(1 - \cos(2\theta)) \\ &= 2\cos^2\theta - 1 \qquad \text{so} \quad \cos^2\theta = \tfrac{1}{2}(1 + \cos(2\theta))\end{aligned}$$

$$\tan(2\theta) = \frac{2\tan\theta}{1 - \tan^2\theta}$$

Sum to product identities

$$\sin A + \sin B = 2\sin\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right) \qquad \cos A + \cos B = 2\cos\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right)$$

$$\sin A - \sin B = 2\cos\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right) \qquad \cos A - \cos B = -2\sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)$$

Product to sum identities

$$\sin A\sin B = \tfrac{1}{2}\cos(A-B) - \tfrac{1}{2}\cos(A+B)$$

$$\cos A\cos B = \tfrac{1}{2}\cos(A+B) + \tfrac{1}{2}\cos(A-B)$$

$$\sin A\cos B = \tfrac{1}{2}\sin(A+B) + \tfrac{1}{2}\sin(A-B)$$

Hyperbolic identities

sinh, cosh and tanh

$$\sinh x = \tfrac{1}{2}\left(e^x - e^{-x}\right) \qquad \cosh x = \tfrac{1}{2}\left(e^x + e^{-x}\right) \qquad \tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

tanh, cosech, sech, coth in terms of sinh and cosh

$$\tanh x = \frac{\sinh x}{\cosh x} \qquad \operatorname{cosech} x = \frac{1}{\sinh x} \qquad \operatorname{sech} x = \frac{1}{\cosh x} \qquad \coth x = \frac{\cosh x}{\sinh x}$$

Hyperbolic Pythagorean identities

$$\cosh^2 x - \sinh^2 x = 1 \qquad 1 - \tanh^2 x = \operatorname{sech}^2 x \qquad \coth^2 x - 1 = \operatorname{cosech}^2 x$$

Symmetry identities

$$\sinh(-x) = -\sinh x \qquad \cosh(-x) = \cosh x \qquad \tanh(-x) = -\tanh x$$

Sum and difference identities

$$\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y$$
$$\sinh(x-y) = \sinh x \cosh y - \cosh x \sinh y$$
$$\tanh(x+y) = \frac{\tanh x + \tanh y}{1 + \tanh x \tanh y}$$
$$\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y$$
$$\cosh(x-y) = \cosh x \cosh y - \sinh x \sinh y$$
$$\tanh(x-y) = \frac{\tanh x - \tanh y}{1 - \tanh x \tanh y}$$

Double-variable and half-variable identities

$$\sinh(2x) = 2\sinh x \cosh x$$

$$\begin{aligned}\cosh(2x) &= \cosh^2 x + \sinh^2 x \\ &= 1 + 2\sinh^2 x \qquad \text{so} \quad \sinh^2 x = \tfrac{1}{2}(\cosh(2x) - 1) \\ &= 2\cosh^2 x - 1 \qquad \text{so} \quad \cosh^2 x = \tfrac{1}{2}(\cosh(2x) + 1)\end{aligned}$$

$$\tanh(2x) = \frac{2\tanh x}{1 + \tanh^2 x}$$

Inverse hyperbolic functions

$$\sinh^{-1} x = \ln\left(x + \sqrt{x^2+1}\right) \qquad (x \in \mathbb{R})$$

$$\cosh^{-1} x = \ln\left(x + \sqrt{x^2-1}\right) \qquad (x \geq 1)$$

$$\tanh^{-1} x = \tfrac{1}{2}\ln\left(\frac{1+x}{1-x}\right) \qquad (-1 < x < 1)$$

Index laws and logarithm laws

Index laws

$a^m a^n = a^{m+n}$ $\qquad \dfrac{a^m}{a^n} = a^{m-n}$ $\qquad (a^m)^n = a^{mn}$

$a^0 = 1$ $\qquad a^{-n} = \dfrac{1}{a^n}$

$(ab)^n = a^n b^n$ $\qquad \left(\dfrac{a}{b}\right)^n = \dfrac{a^n}{b^n}$

$a^{1/n} = \sqrt[n]{a}$ $\qquad a^{m/n} = (\sqrt[n]{a})^m = \sqrt[n]{a^m}$

$\sqrt{ab} = \sqrt{a}\sqrt{b}$ $\qquad \sqrt{\dfrac{a}{b}} = \dfrac{\sqrt{a}}{\sqrt{b}}$

Logarithm laws

$y = \log_b x$ is equivalent to $x = b^y$

$\log_b 1 = 0$ $\qquad \log_b(b^x) = x$

$\log_b b = 1$ $\qquad b^{\log_b x} = x$

$\log_b x + \log_b y = \log_b(xy)$

$\log_b x - \log_b y = \log_b\left(\dfrac{x}{y}\right)$

$r \log_b x = \log_b(x^r)$

Logarithm laws for natural logarithms

$y = \ln x$ is equivalent to $x = e^y$

$\ln 1 = 0$ $\qquad \ln e^x = x$

$\ln e = 1$ $\qquad e^{\ln x} = x$

$\ln x + \ln y = \ln(xy)$

$\ln x - \ln y = \ln\left(\dfrac{x}{y}\right)$

$r \ln x = \ln(x^r)$

Standard derivatives and standard integrals

Standard derivatives

Function	Derivative
a (constant)	0
x^n	nx^{n-1}
e^x	e^x
$\ln x$	$\frac{1}{x}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\sec^2 x$
$\operatorname{cosec} x$	$-\operatorname{cosec} x \cot x$
$\sec x$	$\sec x \tan x$
$\cot x$	$-\operatorname{cosec}^2 x$
$\sin^{-1} x$	$\frac{1}{\sqrt{1-x^2}}$
$\cos^{-1} x$	$-\frac{1}{\sqrt{1-x^2}}$
$\tan^{-1} x$	$\frac{1}{1+x^2}$
$\sinh x$	$\cosh x$
$\cosh x$	$\sinh x$
$\tanh x$	$\operatorname{sech}^2 x$

Standard indefinite integrals

Function	Indefinite integral
a (constant)	$ax + c$
x^n $(n \neq -1)$	$\frac{1}{n+1}\, x^{n+1} + c$
$\frac{1}{x}$	$\ln \lvert x \rvert + c$ or $\ln x + c$, for $x > 0$
e^x	$e^x + c$
$\sin x$	$-\cos x + c$
$\cos x$	$\sin x + c$
$\tan x$	$-\ln \lvert \cos x \rvert + c$
$\sec^2 x$	$\tan x + c$
$\operatorname{cosec}^2 x$	$-\cot x + c$
$\sec x \tan x$	$\sec x + c$
$\operatorname{cosec} x \cot x$	$-\operatorname{cosec} x + c$
$\frac{1}{\sqrt{1-x^2}}$	$\sin^{-1} x + c$ or $-\cos^{-1} x + c$
$\frac{1}{1+x^2}$	$\tan^{-1} x + c$
$\sinh x$	$\cosh x + c$
$\cosh x$	$\sinh x + c$
$\tanh x$	$\ln(\cosh x) + c$
$\operatorname{sech}^2 x$	$\tanh x + c$

Standard series

Standard Taylor series

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \frac{1}{9!}x^9 - \cdots \qquad \text{for } x \in \mathbb{R}$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \frac{1}{8!}x^8 - \cdots \qquad \text{for } x \in \mathbb{R}$$

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots \qquad \text{for } x \in \mathbb{R}$$

$$\ln(1+x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 + \tfrac{1}{5}x^5 - \cdots \qquad \text{for } -1 < x < 1$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots \qquad \text{for } -1 < x < 1$$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots$$

$$(\alpha \text{ can be any real number}) \qquad \text{for } -1 < x < 1$$

Arithmetic series

The finite arithmetic series with first term a, common difference d and n terms has sum

$$a + (a+d) + (a+2d) + \cdots + (a+(n-1)d) = \tfrac{1}{2}n\,(2a+(n-1)d)\,.$$

If the last term of the series is l, the sum is $\frac{1}{2}n(a+l)$.

Geometric series

The finite geometric series with first term a, common ratio $r \neq 1$ and n terms has sum

$$a + ar + ar^2 + \cdots + ar^{n-1} = \frac{a(1-r^n)}{1-r}.$$

The infinite geometric series with first term a and common ratio r, with $-1 < r < 1$, has sum

$$a + ar + ar^2 + \cdots = \frac{a}{1-r}.$$

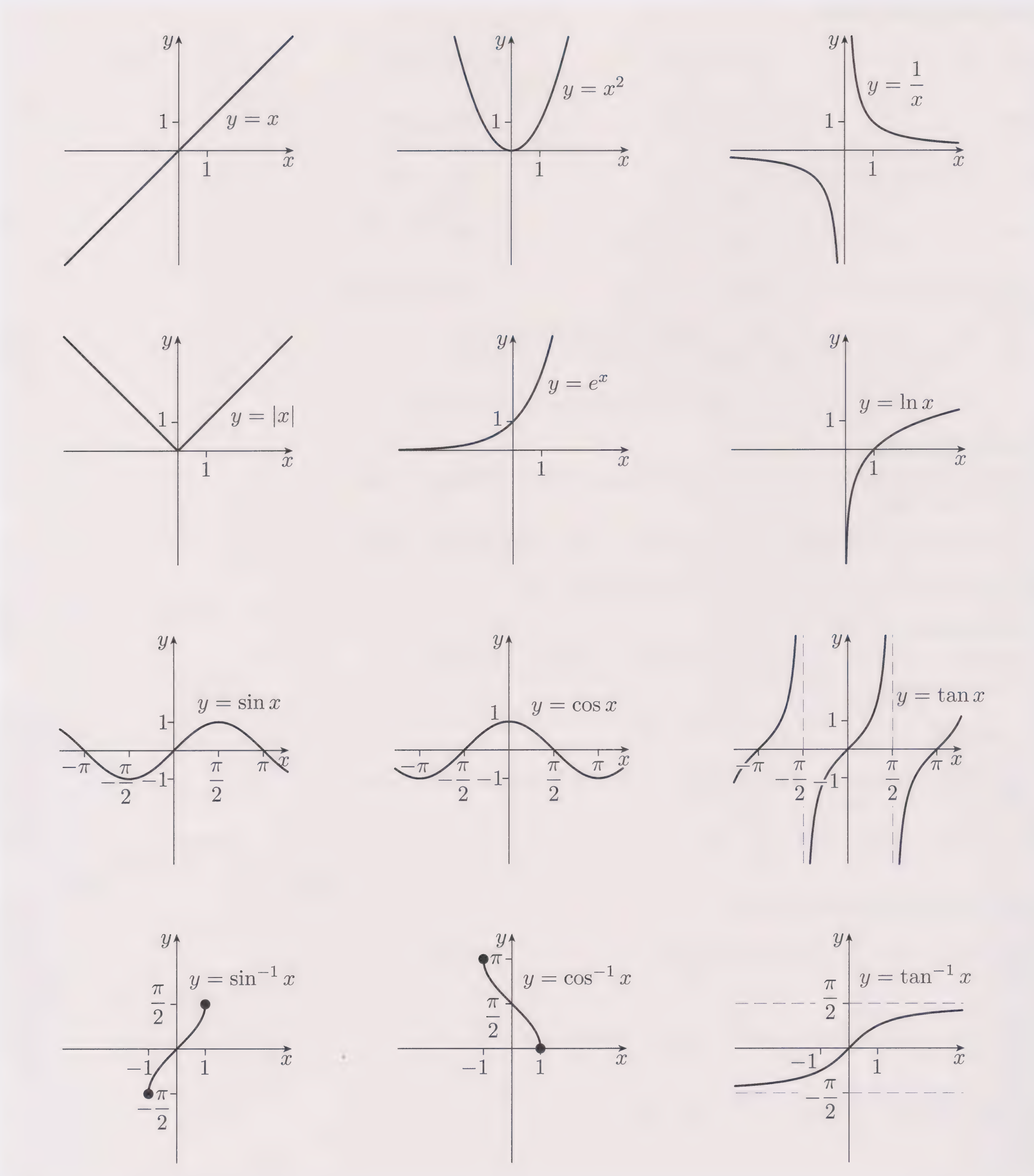
Sums of standard finite series

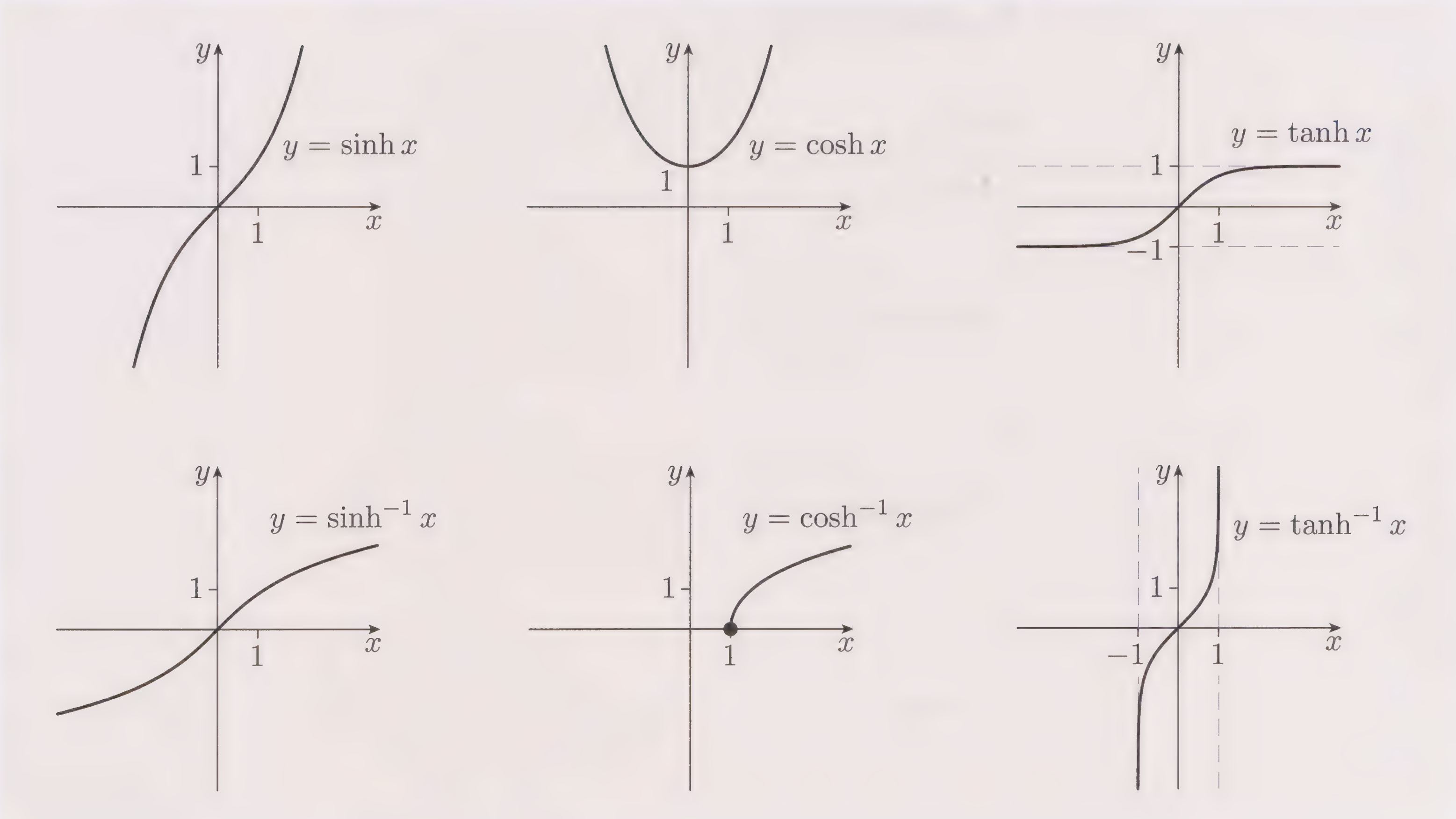
$$\sum_{k=1}^{n} k = 1 + 2 + 3 + \cdots + n = \tfrac{1}{2}n(n+1)$$

$$\sum_{k=1}^{n} k^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2 = \tfrac{1}{6}n(n+1)(2n+1)$$

$$\sum_{k=1}^{n} k^3 = 1^3 + 2^3 + 3^3 + \cdots + n^3 = \tfrac{1}{4}n^2(n+1)^2$$

Useful graphs

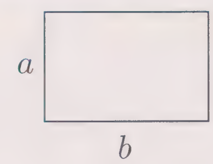
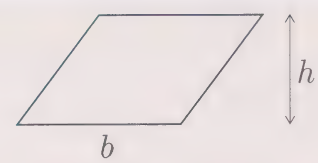
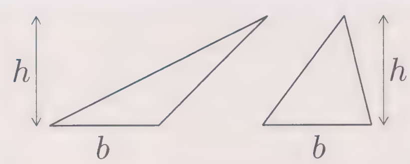
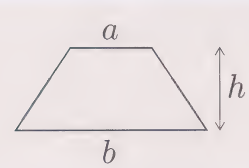




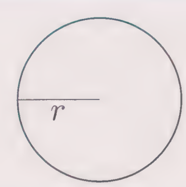
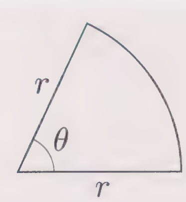
y = sinh x
y = cosh x
y = tanh x
y = sinh⁻¹ x
y = cosh⁻¹ x
y = tanh⁻¹ x

Geometry

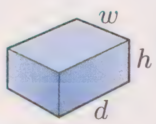
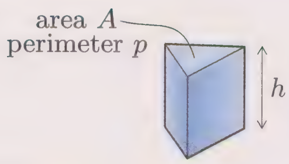
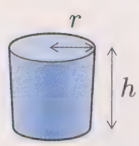
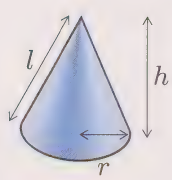
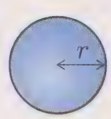
Areas

Shape		Area
Rectangle		ab
Parallelogram		bh
Triangle		$\frac{1}{2}bh$
		$\frac{1}{2}ab\sin\theta$
Trapezium		$\frac{1}{2}(a+b)h$

Circles

Shape		Formulas
Circle		circumference $= 2\pi r$ area $= \pi r^2$
Sector (θ in radians)		arc length $= r\theta$ area $= \frac{1}{2}r^2\theta$

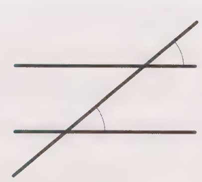
Volumes and surface areas

Solid		Volume	Surface area
Cuboid		whd	$2wh + 2wd + 2hd$
Prism		Ah	$2A + hp$
Cylinder		$\pi r^2 h$	$2\pi r^2 + 2\pi rh$
Cone		$\frac{1}{3}\pi r^2 h$	$\pi r^2 + \pi rl$
Sphere		$\frac{4}{3}\pi r^3$	$4\pi r^2$

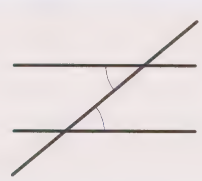
Opposite, corresponding and alternate angles

Where two lines cross,
opposite angles are equal.

Where a line crosses parallel lines,
corresponding angles are equal.



Where a line crosses parallel lines,
alternate angles are equal.



Greek alphabet

A	α	alpha	N	ν	nu
B	β	beta	Ξ	ξ	xi
Γ	γ	gamma	O	o	omicron
Δ	δ	delta	Π	π	pi
E	ε	epsilon	P	ρ	rho
Z	ζ	zeta	Σ	σ	sigma
H	η	eta	T	τ	tau
Θ	θ	theta	Υ	υ	upsilon
I	ι	iota	Φ	ϕ	phi
K	κ	kappa	X	χ	chi
Λ	λ	lambda	Ψ	ψ	psi
M	μ	mu	Ω	ω	omega

SI units

The Système International d'Unités (SI units) is an internationally agreed set of units for measuring physical quantities.

The seven base units are as follows.

Unit	Symbol	Measurement of
metre	m	length
second	s	time
kilogram	kg	mass
kelvin	K	thermodynamic temperature
ampere	A	electric current
candela	cd	luminous intensity
mole	mol	amount of substance

Prefixes may be added to units. The commonly used prefixes are as follows.

Prefix	Symbol	Meaning	Example
nano	n	10^{-9}	nanogram, ng
micro	μ	10^{-6}	microsecond, μs
milli	m	10^{-3}	millisecond, ms
centi	c	10^{-2}	centimetre, cm
kilo	k	10^{3}	kilogram, kg
mega	M	10^{6}	megagram, Mg
giga	G	10^{9}	gigagram, Gg

There are also derived units, which are used for quantities whose measurement combines base units. Some of these are as follows.

Quantity	Unit symbol	Meaning
area	m^2	metres squared or square metres
volume	m^3	metres cubed or cubic metres
speed	$\mathrm{m\,s^{-1}}$	metres per second
acceleration	$\mathrm{m\,s^{-2}}$	metres per second squared

A litre (l) is $0.001\,\mathrm{m}^3$ (or $1000\,\mathrm{cm}^3$). A metric tonne (t) is 1000 kg.

Notation

$\neq$	is not equal to
$\approx$	is approximately equal to
$<$	is less than
$\leq$	is less than or equal to
$>$	is greater than
$\geq$	is greater than or equal to
$\equiv$	is congruent to (modulo n, where n is a positive integer)
$\dots$	ellipsis (dot, dot, dot), used when something has been left out
$\to$	tends to
$\pm$	plus or minus
$\mp$	minus or plus
π	the ratio of the circumference of a circle to its diameter; $\pi = 3.141\,59\dots$
e	the base for natural logarithms; $e = 2.718\,28\dots$
∞	infinity
g	the magnitude of the acceleration due to gravity (taken to be $9.8\,\mathrm{ms}^{-2}$ in MST124 and MST125)
a^n	a to the power n
$\sqrt{a}$	the non-negative square root of the non-negative real number a
$\sqrt[n]{a}$	the non-negative nth root of the non-negative real number a
$\lvert x\rvert$	the modulus (magnitude, absolute value) of the real number x
$n!$	n factorial; $n! = 1 \times 2 \times \cdots \times n$; $0! = 1$
$\mathbb{N}$	the natural numbers: $1, 2, 3, \dots$
$\mathbb{Z}$	the integers: $\dots, -3, -2, -1, 0, 1, 2, 3\dots$
$\mathbb{Q}$	the rational numbers: numbers of the form p/q, where p and q are integers, $q \neq 0$
$\mathbb{R}$	the real numbers
$\mathbb{C}$	the complex numbers
$\mathbb{R}^2$	the Cartesian plane: the set of points (x, y) where x and y are real numbers
$\{a, b, c\}$	the set with elements a, b and c (and similarly for other lists of elements)
$\in$	is a member of, 'is in'
$\notin$	is not a member of, 'is not in'
$\subseteq$	is a subset of
$\cap$	intersection (of sets)
$\cup$	union (of sets)
$\varnothing$	the empty set
$\lvert S\rvert$	the size of the set S
$\overline{S}$	the complement of the set S (in a given set of which S is a subset)
$\angle ABC$	the angle formed by the line segments AB and BC
$\triangle ABC$	the triangle with vertices A, B and C

$f(x)$	the image of the value x under the function f
$\circ$	the symbol for composition of functions; $(g \circ f)(x) = g(f(x))$
f^{-1}	the inverse function of the one-to-one function f
$f(A)$	the image of the set A under the function f
f'	the (first) derivative of the function f
f''	the second derivative of the function f
$f^{(n)}$	the nth derivative of the function f
$\dfrac{\mathrm{d}y}{\mathrm{d}x}$	the (first) derivative of y with respect to x
$\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2}$	the second derivative of y with respect to x
$\dfrac{\mathrm{d}^ny}{\mathrm{d}x^n}$	the nth derivative of y with respect to x
$\dfrac{\mathrm{d}}{\mathrm{d}x}(y)$	a variant of $\dfrac{\mathrm{d}y}{\mathrm{d}x}$
$\dot{x}$	the (first) derivative of x with respect to time t
$\ddot{x}$	the second derivative of x with respect to time t
$\int f(x)\,\mathrm{d}x$	the indefinite integral of $f(x)$ with respect to x
$\int_a^b f(x)\,\mathrm{d}x$	the definite integral of the function f from a to b
$[F(x)]_a^b$	$F(b) - F(a)$
$\mathbf{i}$	the Cartesian unit vector in the direction of the x-axis
$\mathbf{j}$	the Cartesian unit vector in the direction of the y-axis
$\mathbf{k}$	the Cartesian unit vector in the direction of the z-axis
$\mathbf{0}$	the zero vector
$\lvert\mathbf{a}\rvert$	the magnitude (modulus) of the vector $\mathbf{a}$
$\overrightarrow{PQ}$	the displacement vector from P to Q
$\dfrac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$	the (first) derivative of the vector $\mathbf{r}$ with respect to time t
$\dfrac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}$	the second derivative of the vector $\mathbf{r}$ with respect to time t
$\dot{\mathbf{r}}$	the (first) derivative of the vector $\mathbf{r}$ with respect to time t
$\ddot{\mathbf{r}}$	the second derivative of the vector $\mathbf{r}$ with respect to time t

exp	the exponential function
ln	the natural logarithm function; $\log_e$
sin	the sine function
cos	the cosine function
tan	the tangent function
$\sin^{-1}$	the inverse sine function; $\sin^{-1} x$ is the angle in the interval $[-\pi/2, \pi/2]$ whose sine is x
$\cos^{-1}$	the inverse cosine function; $\cos^{-1} x$ is the angle in the interval $[0, \pi]$ whose cosine is x
$\tan^{-1}$	the inverse tangent function; $\tan^{-1} x$ is the angle in the interval $(-\pi/2, \pi/2)$ whose tangent is x
cosec	the cosecant function; $\operatorname{cosec} x = 1/\sin x$
sec	the secant function; $\sec x = 1/\cos x$
cot	the cotangent function; $\cot x = 1/\tan x$
$\sec^{-1}$	the inverse secant function; for any real number x in the set $(-\infty, -1] \cup [1, \infty)$, $\sec^{-1} x$ is the angle in the set $[0, \pi/2) \cup (\pi/2, \pi]$ whose secant is x
sinh	the hyperbolic sine function; $\sinh x = \frac{1}{2}(e^x - e^{-x})$
cosh	the hyperbolic cosine function; $\cosh x = \frac{1}{2}(e^x + e^{-x})$
tanh	the hyperbolic tangent function; $\tanh x = \sinh x/\cosh x$
$\sinh^{-1}$	the inverse hyperbolic sine function; for any real number x, $\sinh^{-1} x$ is the real number whose hyperbolic sine is x
$\cosh^{-1}$	the inverse hyperbolic cosine function; for any number x in the interval $[1, \infty)$, $\cosh^{-1} x$ is the number in the interval $[0, \infty)$ whose hyperbolic cosine is x
$\tanh^{-1}$	the inverse hyperbolic tangent function; for any number x in the interval $(-1, 1)$, $\tanh^{-1} x$ is the real number whose hyperbolic tangent is x
cosech	the hyperbolic cosecant function; $\operatorname{cosech} x = 1/\sinh x$
sech	the hyperbolic secant function; $\operatorname{sech} x = 1/\cosh x$
coth	the hyperbolic cotangent function; $\coth x = 1/\tanh x$
$\sin^2 \theta$	$(\sin \theta)^2$; a similar notation is used for other powers (but not -1) and for other trigonometric and hyperbolic functions

$\mathbf{I}$	an identity matrix
$\mathbf{A}^{-1}$	the inverse of the invertible matrix $\mathbf{A}$
$\det \mathbf{A}$	the determinant of the square matrix $\mathbf{A}$
$\lvert\mathbf{A}\rvert$	the determinant of the square matrix $\mathbf{A}$
$\operatorname{tr} \mathbf{A}$	the trace of the square matrix $\mathbf{A}$; $\operatorname{tr} \mathbf{A}$ is the sum of the elements on the leading diagonal of $\mathbf{A}$
(x_n)	the infinite sequence $x_1, x_2, x_3, \ldots$
$(x_n)_{n=p}^{q}$	the finite sequence $x_p, x_{p+1}, x_{p+2}, \ldots, x_q$
$(x_n)_{n=p}^{\infty}$	the infinite sequence $x_p, x_{p+1}, x_{p+2}, \ldots$
$\displaystyle\sum_{n=p}^{q} x_n$	the finite sum $x_p + x_{p+1} + x_{p+2} + \cdots + x_q$
$\displaystyle\sum_{n=p}^{\infty} x_n$	the infinite sum $x_p + x_{p+1} + x_{p+2} + \cdots$
${}^nC_k$	the number of combinations of n objects taken k at a time; a binomial coefficient; $\displaystyle {}^nC_k = \frac{n!}{k!\,(n-k)!}$
${}^nP_k$	the number of permutations of n objects taken k at a time; $\displaystyle {}^nP_k = \frac{n!}{(n-k)!}$
$P(E)$	the probability of the event E occurring
i	imaginary number whose square is -1
$\operatorname{Re}(z)$	the real part of the complex number z
$\operatorname{Im}(z)$	the imaginary part of the complex number z
$\overline{z}$	the complex conjugate of the complex number z
$\lvert z\rvert$	the modulus of the complex number z
$\operatorname{Arg}(z)$	the principal argument of the complex number z
$\exists$	there exists
$\forall$	for all
$\Rightarrow$	implies
$\Leftrightarrow$	if and only if

MST124

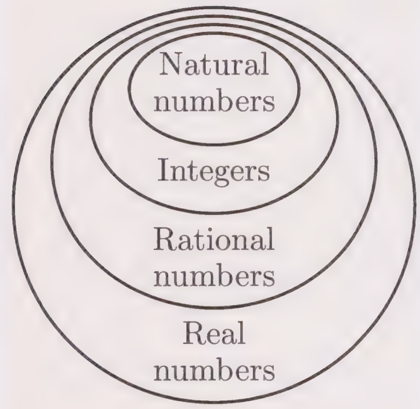
MST124 Unit 1 Algebra

Numbers

Types of numbers

The **natural numbers** (or **positive integers**) are $1,\ 2,\ 3,\ \ldots$

The **integers** are $\ldots,\ -3,\ -2,\ -1,\ 0,\ 1,\ 2,\ 3,\ \ldots$



The **rational numbers** are the numbers that can be written in the form integer/integer, or, equivalently, the decimal numbers that terminate or recur.

The **irrational numbers** are the numbers that cannot be written in the form integer/integer, or, equivalently, the decimal numbers with an infinite number of digits after the decimal point but with no block of digits that repeats indefinitely.

The **real numbers** are the rational numbers together with the irrational numbers. Each real number corresponds to a point on the number line.

A **prime number** (or **prime**) is an integer greater than 1 whose only positive factors are 1 and itself.

A **composite number** is an integer greater than 1 that is not a prime number.

A **square number** is an integer that can be written as the square of an integer.

The prime numbers under 100

2 3 5 7 11 13 17 19 23 29 31 37 41 43 47 53 59 61 67 71 73 79 83 89 97

The square numbers up to 15^2

0 1 4 9 16 25 36 49 64 81 100 121 144 169 196 225

The fundamental theorem of arithmetic

Every integer greater than 1 can be written as a product of prime factors in just one way (except that you can change the order of the factors).

Factors and multiples of integers

A **factor** of an integer is an integer that divides exactly into the first integer. A **multiple** of an integer is an integer into which the first integer divides exactly.

A **factor pair** of an integer is a pair of its factors that multiply together to give the integer. A **positive factor pair** of a positive integer is a pair of its positive factors that multiply together to give the integer.

The **prime factorisation** of an integer greater than 1 is any expression that shows it written as a product of prime factors.

Common factors and common multiples of integers

A **common multiple** of two or more integers is an integer that is a multiple of all of them. The **lowest** (or **least**) **common multiple** (**LCM**) of two or more integers is the smallest positive integer that is a multiple of all of them.

A **common factor** of two or more integers is an integer that is a factor of all of them. The **highest common factor** (**HCF**)(or **greatest common divisor (GCD)**) of two or more integers is the largest positive integer that is a factor of all of them.

To find the lowest common multiple or highest common factor of two or more integers greater than 1

- Find the prime factorisations of the numbers.
- To find the LCM, multiply together the highest power of each prime factor occurring in any of the numbers.
- To find the HCF, multiply together the lowest power of each prime factor common to all the numbers.

Powers and roots

The nth **power** of a number a, denoted by a^n, is obtained by multiplying together n copies of the number. Here a is the **base number** or **base**, and n is the **power**, **index** or **exponent**.

An nth **root** of a number is a number that when raised to the nth power gives the original number. A **square root** is a second root and a **cube root** is a third root. The non-negative square root of a non-negative number a is denoted by $\sqrt{a}$. The non-negative nth root of a non-negative number a is denoted by $\sqrt[n]{a}$.

The **reciprocal** of a number a is $1/a$, which can also be written as a^{-1}.

To express a number in **scientific notation**, write it in the form

(a number between 1 and 10, but not including 10)
$\times$ (an integer power of ten).

For the index laws, see page 7.

Surds

A **surd** is a numerical expression (such as $1 - 2\sqrt{5}$) that contains one or more irrational roots of numbers.

To manipulate surds, use the usual rules of algebra and the index laws.

To **rationalise** the denominator of a surd that has an irrational denominator, multiply the numerator and denominator by a suitable surd. If the denominator is a sum of two terms, either or both of which is a rational number multiplied by an irrational square root, then a suitable surd is the expression obtained by changing the sign of one of the two terms in the sum. This expression is called a **conjugate** of the expression in the denominator. (For example, a conjugate of $\sqrt{2} - 3\sqrt{5}$ is $\sqrt{2} + 3\sqrt{5}$.)

Expressions

An **expression** is an arrangement of letters, numbers and/or mathematical symbols, which is such that if values are substituted for any letters present, then you can work out the value of the arrangement.

An **algebraic expression** includes letters. Each such letter may be:

- a **variable**, representing any number, or any number of a particular kind
- an **unknown**, representing a particular number that you do not know, but usually you want to discover
- a **constant**, representing a particular number whose value is fixed.

If an expression is a list of quantities that are added, then these quantities are called the **terms** of the expression.

If a term consists of a number multiplied by powers of variables, then the number is called the **coefficient** of the term.

An **algebraic fraction** is an algebraic expression written in the form of a fraction. The **numerator** and **denominator** of the fraction are its top and bottom, respectively.

Factors and multiples of algebraic expressions

Roughly speaking, if an algebraic expression can be written in the form

$$\text{something} \times \text{something},$$

then both 'somethings' are **factors** of the expression, and the expression is a **multiple** of both 'somethings'.

A **common factor** of two or more algebraic expressions is an expression that is a factor of all of them.

A **common multiple** of two or more algebraic expressions is an expression that is a multiple of all of them.

A **highest common factor** of two or more algebraic expressions is a common factor that is a multiple of all other common factors.

A **lowest common multiple** of two or more algebraic expressions is a common multiple that is a factor of all other common multiples.

Factorising an expression means writing it as the product of two or more expressions, neither of which is 1 (and, usually, neither of which is -1).

Difference of two squares

$$(A+B)(A-B) = A^2 - B^2$$

Squaring brackets

$$(A+B)^2 = A^2 + 2AB + B^2$$
$$(A-B)^2 = A^2 - 2AB + B^2$$

Equations

An **equation** consists of two expressions, with an equals sign between them.

The **solutions** of an equation are the values of its variables that make the equation true. These values **satisfy** the equation.

An **identity** is an equation that is satisfied by all possible values of its variables.

A **linear equation** is an equation in which, after you have expanded any brackets and cleared any fractions, each term is either a constant term or a constant value times a variable.

If an equation contains more than one variable, and one side of the equation is just a single variable that does not appear at all on the other side, then that variable is called the **subject** of the equation.

To rearrange an equation

Carry out any of the following operations on an equation to obtain an equivalent equation.

- Rearrange the expressions on one or both sides.
- Swap the sides.
- Do any of the following things to both sides:
  - add or subtract something
  - multiply or divide by something (provided that it is non-zero)
  - raise to a power (provided that the power is non-zero, and that the expressions on each side of the equation can take only non-negative values).

To make a variable the subject of an equation

(This works for some equations but not all.)

Use the rules for rearranging equations to obtain successive equivalent equations. Aim to obtain an equation in which the required subject is alone on one side. To achieve this, do the following, in order.

1. Clear any fractions and multiply out any brackets. To clear fractions, multiply through by a suitable expression.
2. Add or subtract terms on both sides to get all the terms containing the required subject on one side, and all the other terms on the other side. Collect like terms.
3. If more than one term contains the required subject, then take it out as a common factor.
4. Divide both sides by the expression that multiplies the required subject.

MST124 Unit 2 Graphs and equations

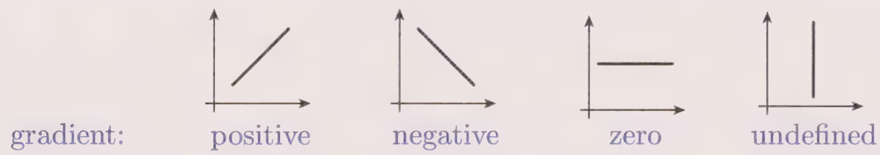
Mathematical modelling

A **mathematical model** is a collection of assumptions and mathematical statements that is intended to describe how some phenomenon in the real world behaves, and to enable predictions to be made about its behaviour.

A **linear mathematical model** is a mathematical model based on a relationship between two variables that is represented by a straight line.

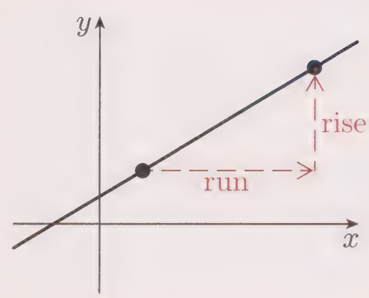
Straight-line graphs

The **gradient** of a straight line is the number of units moved up for every one unit moved to the right.



The gradient of the line through the points (x_1, y_1) and (x_2, y_2), where $x_1 \neq x_2$, is given by

$$\text{gradient} = \frac{\text{rise}}{\text{run}} = \frac{y_2 - y_1}{x_2 - x_1}.$$



The x- and y-**intercepts** of a straight line are the values of x and y where the line crosses the x-axis and the y-axis, respectively.

Equations of lines

The straight line with gradient m and y-intercept c has equation

$$y = mx + c.$$

The horizontal line with y-intercept c has equation $y = c$.

The vertical line with x-intercept d has equation $x = d$.

The straight line with gradient m that passes through the point (x_1, y_1) has equation

$$y - y_1 = m(x - x_1).$$

Parallel and perpendicular lines

Two straight lines are **parallel** if they never cross, even when extended infinitely far in each direction. The gradients of any two parallel lines are equal (or both lines are vertical).

Two straight lines are **perpendicular** if they are at right angles to each other. The gradients of any two perpendicular lines (not parallel to the axes) have product -1.

Simultaneous linear equations

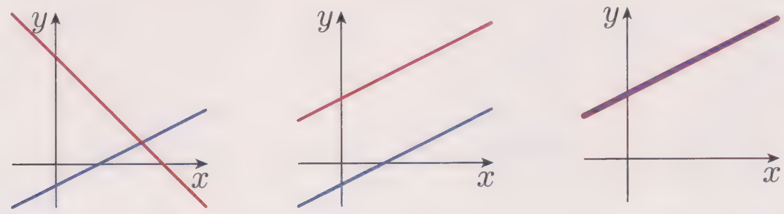
A pair of **simultaneous linear equations** consists of two equations of the form

$$ax + by = e$$
$$cx + dy = f$$

that apply simultaneously. Here x and y are unknowns, and a, b, c, d, e and f are constants.

A **solution** of a pair of simultaneous equations is a pair of values of the unknowns x and y that satisfy both equations. These values are the coordinates of the point of intersection of the lines represented by the equations.

A pair of simultaneous linear equations may have one solution, no solutions, or infinitely many solutions.



One solution No solutions Infinitely many solutions

To solve a pair of simultaneous equations: substitution method

1. Rearrange one of the equations, if necessary, to obtain a formula for one unknown in terms of the other.
2. Use this formula to substitute for this unknown in the other equation.
3. You now have an equation in one unknown. Solve it to find the value of that unknown.
4. Substitute this value into an equation involving both unknowns to find the value of the other unknown.

(Check: confirm that the two values satisfy the original equations.)

To solve a pair of simultaneous equations: elimination method

1. Multiply one or both of the equations by suitable numbers, if necessary, to obtain two equations that can be added or subtracted to eliminate one of the unknowns.
2. Add or subtract the equations to eliminate this unknown.
3. You now have an equation in one unknown. Solve it to find the value of that unknown.
4. Substitute this value into an equation involving both unknowns to find the value of the other unknown.

(Check: confirm that the two values satisfy the original equations.)

Quadratic expressions

A **quadratic expression**, or **quadratic**, is an expression of the form $ax^2 + bx + c$, where a, b and c are constants with $a \neq 0$.

To factorise a quadratic of the form $x^2 + bx + c$

1. Start by writing $x^2 + bx + c = (x \quad)(x \quad)$.
2. Find the factor pairs of c (including both positive and negative ones).
3. Choose the factor pair with sum b, if there is such a pair.
4. Write your factor pair p, q in position: $x^2 + bx + c = (x + p)(x + q)$.

To factorise a quadratic of the form $ax^2 + bx + c$

1. Take out any numerical common factors. If the coefficient of x^2 is negative, also take out the factor -1. Then apply the steps below to the quadratic inside the brackets.
2. Find the positive factor pairs of a, the coefficient of x^2. For each such factor pair d, e write down a framework $(dx \quad)(ex \quad)$.
3. Find all the factor pairs of c, the constant term (including both positive and negative ones).
4. For each framework and each factor pair of c, write the factor pair in the gaps in the framework in both possible ways.
5. For each of the resulting cases, calculate the term in x that you obtain when you multiply out the brackets.
6. Identify the case where this term is bx, if there is such a case. This is the required factorisation.

To complete the square in a quadratic of the form $x^2 + bx$

1. Write down $(x \quad)^2$, filling the gap with the number that is half of b, the coefficient of x (including its $+$ or $-$ sign).
2. Subtract the square of the number that you wrote in the gap.

To complete the square in a quadratic of the form $x^2 + bx + c$

1. Use the strategy above to complete the square in the subexpression $x^2 + bx$.
2. Collect the constant terms.

To complete the square in a quadratic of the form $ax^2 + bx + c$, where $a \neq 1$

1. Rewrite the quadratic with the coefficient a taken out of the subexpression $ax^2 + bx$ as a factor. This generates a pair of brackets.
2. Use the earlier strategy to complete the square in the simple quadratic inside the brackets. This generates a second pair of brackets, inside the first pair.
3. Multiply out the outer brackets.
4. Collect the constant terms.

Quadratic equations

A **quadratic equation** is an equation of the form $ax^2 + bx + c = 0$, where a, b and c are constants with $a \neq 0$. It has at most two solutions.

To simplify a quadratic equation

- If necessary, rearrange the equation so that all the non-zero terms are on the same side.
- If the coefficient of x^2 is negative, then multiply the equation through by -1 to make this coefficient positive.
- If the coefficients have a common factor, then divide the equation through by this factor.
- If any of the coefficients are fractions, then multiply the equation through by a suitable number to clear them.

To solve a quadratic equation $ax^2 + bx + c = 0$

Simplify it, then use one of the following three methods.

Factorisation

1. Factorise the quadratic expression.
2. Use the fact that if the product of two numbers is zero then at least one of the numbers must be zero.
3. Solve the resulting two linear equations.

Completing the square

1. Complete the square in the quadratic expression.
2. Rearrange the equation to obtain the square alone on the left-hand side and a constant on the right-hand side.
3. Take square roots of both sides (remembering that a positive number has both a positive and a negative square root).
4. Solve the resulting two linear equations.

The quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

The number of real solutions of a quadratic equation

The **discriminant** of the quadratic expression $ax^2 + bx + c$ is the value $b^2 - 4ac$.

The quadratic equation $ax^2 + bx + c = 0$ (where $a \neq 0$) has:

- two real solutions if $b^2 - 4ac > 0$ (the discriminant is positive)
- one real solution if $b^2 - 4ac = 0$ (the discriminant is zero)
- no real solutions if $b^2 - 4ac < 0$ (the discriminant is negative).

A **repeated solution** of a quadratic equation is a solution that is the only solution of the equation.

Quadratic graphs

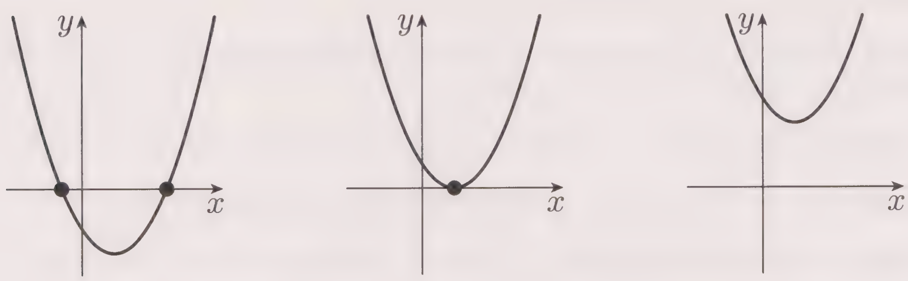
The shape of the graph of an equation of the form $y = ax^2 + bx + c$, where a, b and c are constants with $a \neq 0$, is called a **parabola**.

a positive: u-shaped parabola

a negative: n-shaped parabola

Properties of the graph of $y = ax^2 + bx + c$, where $a \neq 0$

- The graph is a parabola with a vertical axis of symmetry.
- If a is positive it is **u-shaped**; if a is negative it is **n-shaped**.
- It has two, one or no x-intercepts .
- It has one y-intercept.



u-shaped parabolas with two, one and no x-intercepts

To sketch the graph of $y = ax^2 + bx + c$, where $a \neq 0$

1. Find whether the parabola is u-shaped or n-shaped.
2. Find its intercepts, axis of symmetry and vertex.
3. Plot the features found, and hence sketch the parabola.
4. Label the parabola with its equation, its intercepts and the coordinates of its vertex.

Sketching versus plotting

A **sketch** of a graph is a diagram that gives an impression of its shape, with key points marked and positioned approximately correctly relative to a pair of coordinate axes.

A **plot** of a graph is a more accurate diagram obtained by precisely plotting a reasonably large number of points on the graph.

Displacement and velocity along a straight line

For an object moving along a straight line,

- its **displacement** from a reference point is its distance from the reference point, with a plus or minus sign to indicate the direction
- its **velocity** is its speed, with a plus or minus sign to indicate the direction.

MST124 Unit 3 Functions

Sets

A **set** is a collection of objects.

The notation $x \in A$ means that the object x is in the set A.

If A and B are any two sets, then

- their **intersection** $A \cap B$ is the set whose members are all the elements that belong to both A and B
- their **union** $A \cup B$ is the set whose members are all the elements that belong to either A or B (or both).

These definitions extend to intersections and unions of more than two sets.

A **subset** of a set A is a set whose elements all belong to A.

The **empty set**, denoted by $\varnothing$, is the set that contains no elements.

The set containing the single element a can be denoted by $\{a\}$. Similarly, the set containing the elements a and b can be denoted by $\{a, b\}$, and so on.

Intervals

An **interval** is a set of real numbers that corresponds to a part of the number line that you can draw 'without lifting your pen from the paper'. A number that lies at an end of an interval is called an **endpoint**.

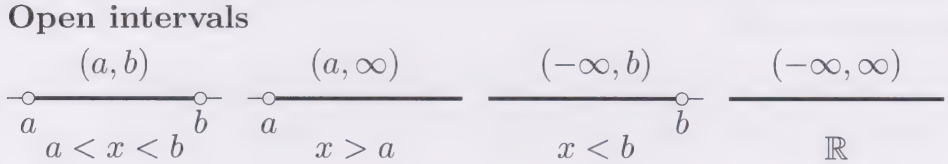
A **closed** interval includes all of its endpoints.

An **open** interval does not include any of its endpoints.

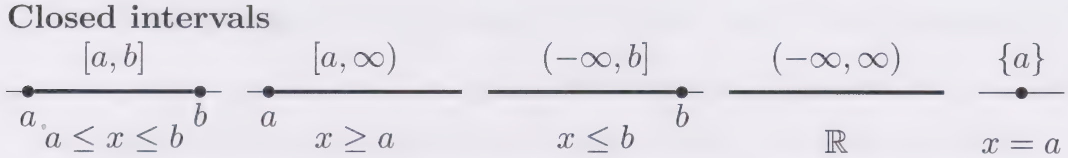
A **half-open** (or **half-closed**) interval includes one endpoint and excludes another.

The interval $\mathbb{R}$ has no endpoints, so it is both open and closed.

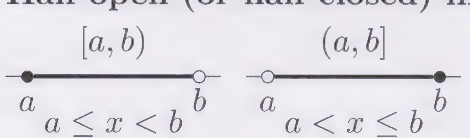
Open intervals



Closed intervals



Half-open (or half-closed) intervals



Functions

A **function** consists of:

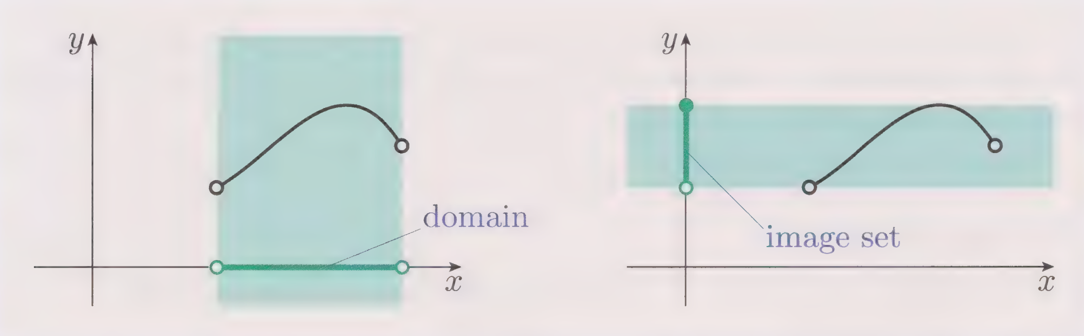
- a set of allowed input values, called the **domain** of the function
- a set of values in which every output value lies, called the **codomain** of the function
- a process, called the **rule** of the function, for converting each input value into exactly one output value.

If f is a function and x is a value in its domain, then the **image of** x **under** f, or the **value of** f **at** x, denoted by $f(x)$, is the output value corresponding to the input value x.

The **image set** of a function is the set consisting of all the values in its codomain that occur as output values.

A **real function** is a function whose domain and codomain are both subsets of $\mathbb{R}$. In MST124 the word 'function' is assumed to mean 'real function', and it is assumed that the codomain of every function is $\mathbb{R}$.

The **graph** of a function f consists of the points with coordinates $(x, f(x))$ where x is a value in the domain of f.



When a function is specified using an equation that expresses one variable in terms of another variable, the input variable is called the **independent variable** and the output variable is called the **dependent variable**.

A **piecewise-defined function** is a function whose rule is specified by using different formulas for different parts of its domain.

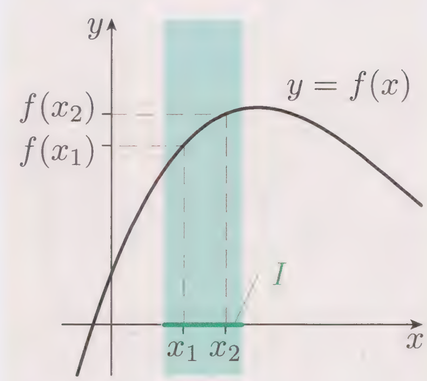
Domain convention

When a function is specified by just a rule, it is understood that the domain of the function is the largest possible set of values for which the rule is applicable.

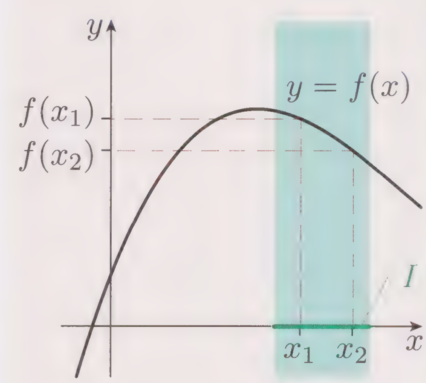
Functions increasing or decreasing on an interval

A function f is **increasing on the interval** I if for all values x_1 and x_2 in I such that $x_1 < x_2$, we have $f(x_1) < f(x_2)$.

A function f is **decreasing on the interval** I if for all values x_1 and x_2 in I such that $x_1 < x_2$, we have $f(x_1) > f(x_2)$.



A function f increasing on an interval I



A function f decreasing on an interval I

Composite functions

The **composite function** $g \circ f$ of the functions f and g is the function whose rule is

$$(g \circ f)(x) = g(f(x)),$$

and whose domain consists of all the values x in the domain of f such that $f(x)$ is in the domain of g.

Inverse functions

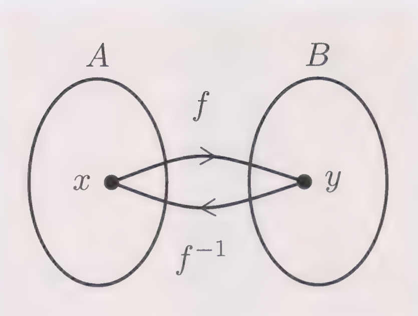
One-to-one functions

A function f is **one-to-one** (or **invertible**) if for all numbers x_1 and x_2 in its domain such that $x_1 \neq x_2$, we have $f(x_1) \neq f(x_2)$. (That is, no two different input values give the same output value.)

Only one-to-one functions have inverse functions.

If a function is either increasing on its whole domain or decreasing on its whole domain, then it is one-to-one.

Inverse functions



If f is a one-to-one function with domain A and image set B, then the **inverse function**, or **inverse**, of f, denoted by f^{-1}, is the function with domain B whose rule is given by

$$f^{-1}(y) = x, \quad \text{where} \quad f(x) = y.$$

The image set of f^{-1} is A.

If a function f has inverse f^{-1}, then the function f^{-1} has inverse f.

For any pair of inverse functions f and f^{-1},

$$(f^{-1} \circ f)(x) = x, \quad \text{for every value } x \text{ in the domain of } f\text{, and}$$
$$(f \circ f^{-1})(x) = x, \quad \text{for every value } x \text{ in the domain of } f^{-1}.$$

To find the rule of the inverse function of a one-to-one function f

1. Write $y = f(x)$ and rearrange this equation to express x in terms of y.
2. Use the resulting equation $x = f^{-1}(y)$ to write down the rule of f^{-1}. (Usually, change the input variable from y to x.)

Graphs of inverse functions

The graphs of every pair of inverse functions are the reflections of each other in the line $y = x$ (when the coordinate axes have equal scales).

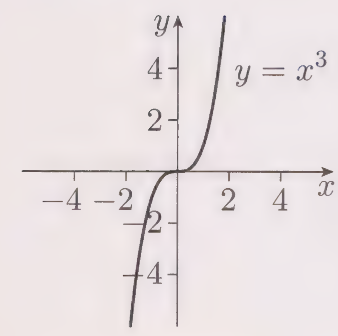
Functions that are not one-to-one

A **restriction** of a function is a function obtained by keeping the rule the same, but removing some values from the domain.

For a function that is not one-to-one and so has no inverse function, we sometimes use the inverse function of a one-to-one restriction of the function.

Some standard functions

Polynomial functions



A **polynomial expression** in x is a sum of finitely many terms, each of the form ax^n where a is a number and n is a non-negative integer.

A **polynomial function** is a function whose rule is of the form '$f(x) =$ a polynomial expression in x'.

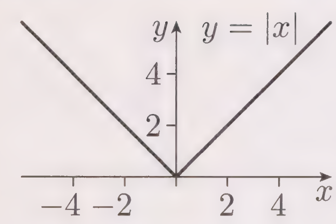
The **degree** of a polynomial expression or function is the highest power of the variable (usually x) in the expression or function.

A **linear function** is a polynomial function of degree at most 1.

A **quadratic**, **cubic**, **quartic** or **quintic** function is a polynomial function of degree 2, 3, 4 or 5, respectively.

The terms **linear**, **quadratic**, **cubic**, **quartic** and **quintic** are used for polynomial expressions in the same way.

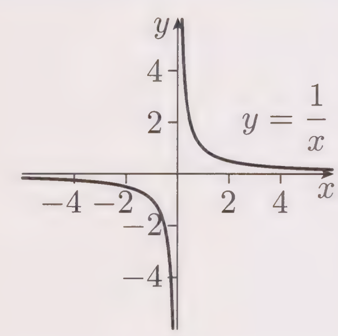
Modulus function



The **modulus**, **magnitude** or **absolute value** $|x|$ of a real number x is its 'distance from zero', or its 'value without its sign'. For example, $|3| = 3$ and $|-3| = 3$.

The **modulus function** is the function $f(x) = |x|$.

Reciprocal function

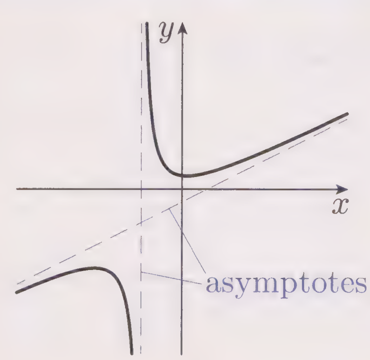


The **reciprocal function** is the function $f(x) = 1/x$. Its domain contains all real numbers except 0.

Rational functions

A **rational function** is a function whose rule is of the form $f(x) = p(x)/q(x)$, where p and q are polynomial functions.

The reciprocal function and every polynomial function are rational functions.



An **asymptote** is a line that a curve approaches arbitrarily closely as the distance along the curve away from the origin increases.

Many rational functions have asymptotes, which can be horizontal, vertical or slant.

Exponential and logarithmic functions

For details of these functions, see page 34.

Trigonometric functions

For details of these functions, see page 39.

Translating, reflecting and scaling graphs

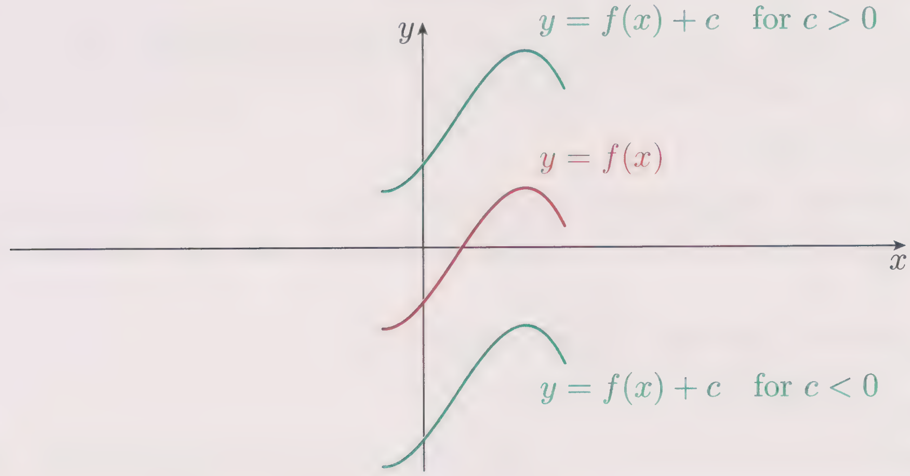
On this page and the next, f is a function and c is a constant.

The scalings and translations on these pages can be applied in any order with the same result, except that the order of a vertical translation and a vertical scaling relative to each other affects the result, and similarly for a horizontal translation and a horizontal scaling.

Translating graphs vertically

To obtain the graph of $y = f(x) + c$,

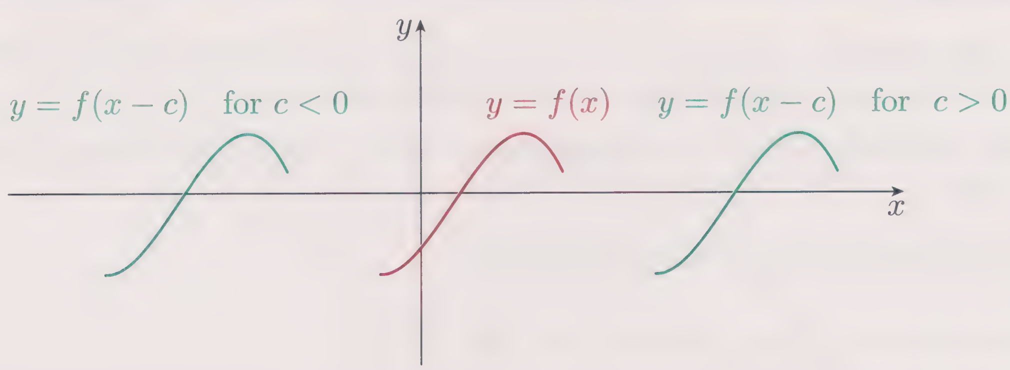
> **translate** the graph of $y = f(x)$ up by c units (the translation is down if c is negative).



Translating graphs horizontally

To obtain the graph of $y = f(x - c)$,

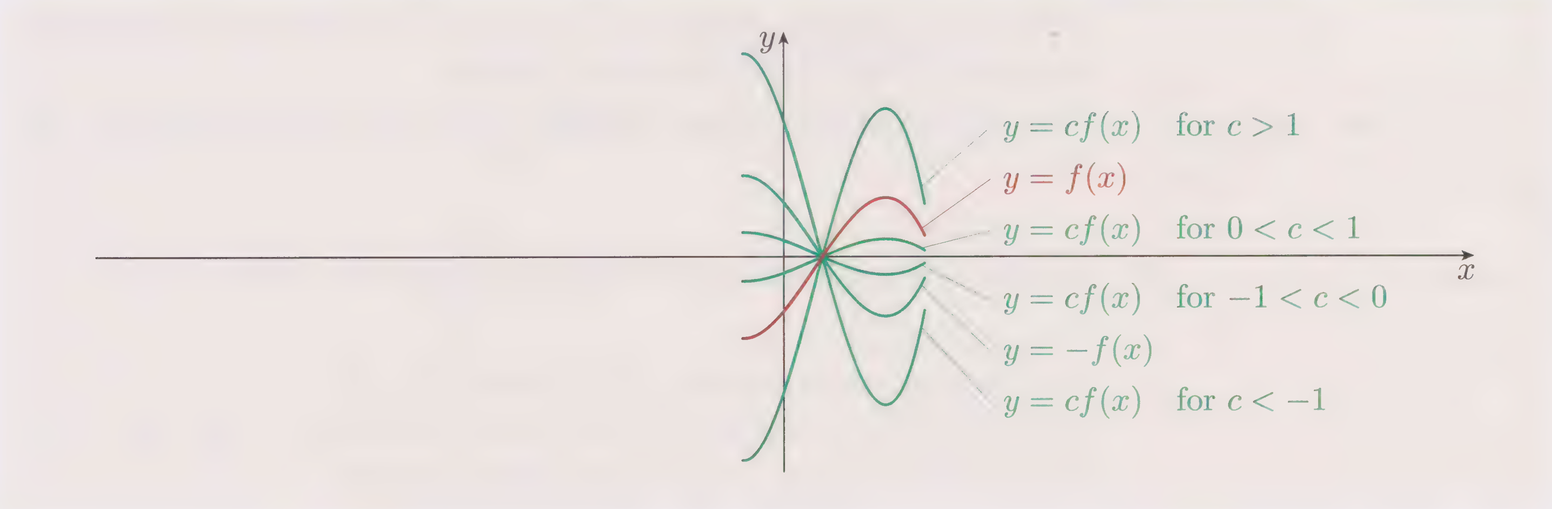
> **translate** the graph of $y = f(x)$ to the right by c units (the translation is to the left if c is negative).



Scaling graphs vertically

To obtain the graph of $y = cf(x)$,

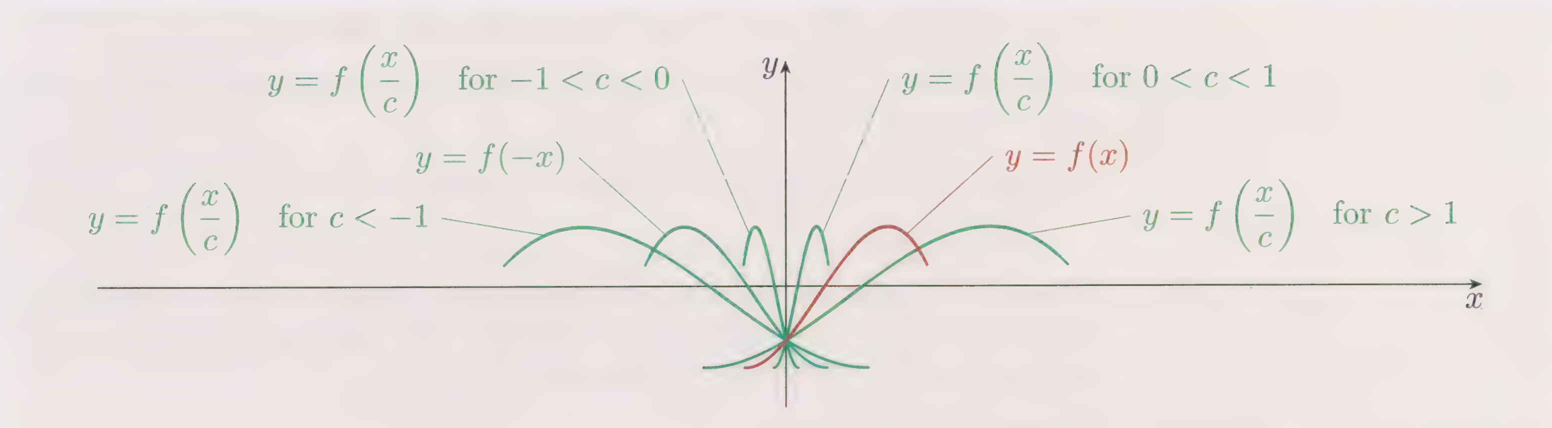
scale the graph of $y = f(x)$ vertically by a factor of c.



Scaling graphs horizontally

To obtain the graph of $y = f\left(\dfrac{x}{c}\right)$,

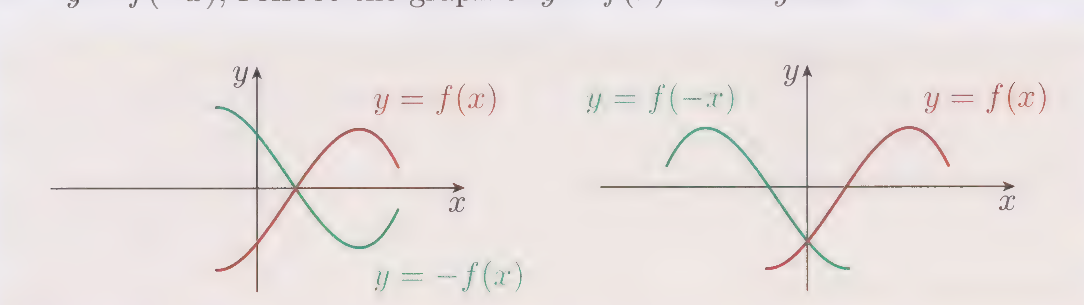
scale the graph of $y = f(x)$ horizontally by a factor of c.

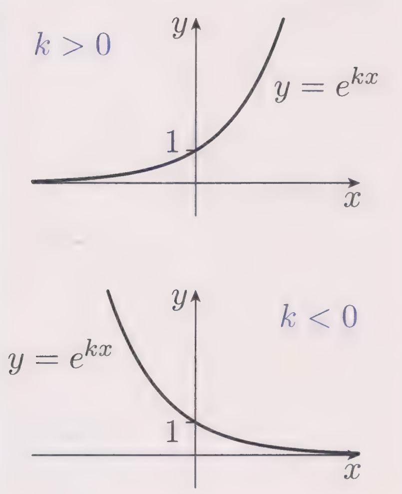


Reflecting graphs in the coordinate axes

To obtain the graph of

- $y = -f(x)$, **reflect** the graph of $y = f(x)$ in the x-axis
- $y = f(-x)$, **reflect** the graph of $y = f(x)$ in the y-axis.





Exponential functions

An **exponential function** is a function whose rule is of the form

$$f(x) = b^x,$$

where b is a positive constant, not equal to 1. The number b is the **base number** or **base** of the exponential function.

Equivalently, an **exponential function** is a function whose rule is of the form

$$f(x) = e^{kx},$$

where k is a non-zero constant (and e is the special constant $2.718\ldots$).

The exponential function

The **exponential function** is the function $f(x) = e^x$.
It has the special property that its gradient is exactly 1 at the point $(0, 1)$.
The expression e^x can be written as $\exp x$ or $\exp(x)$.

Logarithms

The **logarithm to base** b of a number x, denoted by $\log_b x$, is the power to which the base b must be raised to give the number x. So the equations

$$y = \log_b x \quad \text{and} \quad x = b^y$$

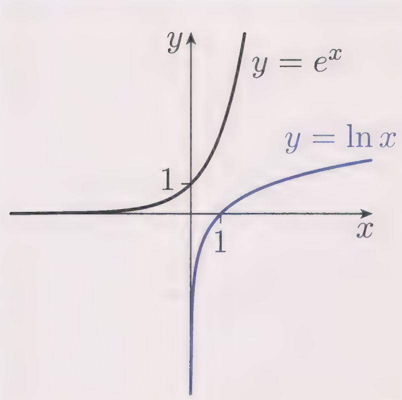
are equivalent. The base b must be positive and not equal to 1.

Only positive numbers have logarithms, but logarithms themselves can be any number.

A **logarithmic function** is a function whose rule is of the form

$$f(x) = \log_b x,$$

where b is a positive constant, not equal to 1. The number b is the **base** or **base number** of the logarithmic function. The functions $f(x) = \log_b x$ and $g(x) = b^x$ are inverses of each other.



Natural logarithms

The **natural logarithm** of a number x, denoted by $\ln x$, is its logarithm to base e. That is, it is the power to which e must be raised to give the number x. So the equations

$$y = \ln x \quad \text{and} \quad x = e^y$$

are equivalent.

The **natural logarithm function** is the function $f(x) = \ln x$.
The functions $f(x) = \ln x$ and $g(x) = e^x$ are inverses of each other.

For the logarithm laws, see page 7.

Exponential equations

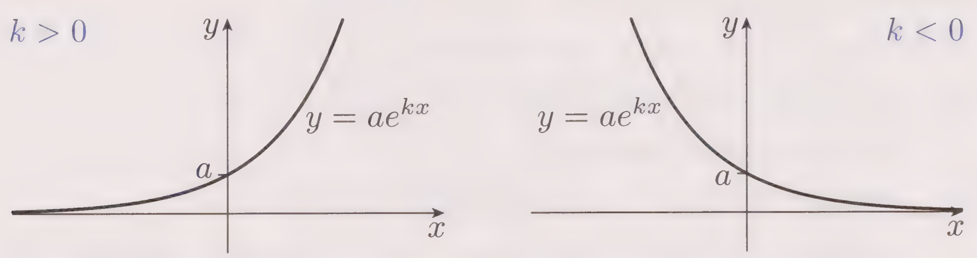
An **exponential equation** is an equation in which an unknown is in an exponent. To solve an exponential equation, take logarithms of both sides.

Exponential growth and decay

A quantity **changes exponentially** if its change can be modelled by a function whose rule is of the form $f(x) = ae^{kx}$, where a and k are non-zero constants.

If a and k are both positive, then the quantity **grows exponentially**, the function is an **exponential growth function**, and the graph is an **exponential growth curve**.

If a is positive but k is negative, then the quantity **decays exponentially**, the function is an **exponential decay function**, and the graph is an **exponential decay curve**.



Exponential growth curve

Exponential decay curve

Growth and decay factors

If $f(x) = ae^{kx}$, then whenever x is increased by p units, the value of $f(x)$ is multiplied by e^{kp}.

If $p > 0$, this factor e^{kp} is called the **growth factor** of f for the period p or the **decay factor** of f for the period p, according to whether f is an exponential growth or decay function. Growth factors are greater than 1, and decay factors are between 0 and 1, exclusive.

Doubling and halving periods

The **doubling period** of the exponential growth function f is the value of p such that whenever x is increased by p the value of $f(x)$ doubles.

The **halving period** of the exponential decay function f is the value p such that whenever x is increased by p the value of $f(x)$ halves.

To find a doubling or halving period

If $f(x) = ae^{kx}$ is an exponential growth function (so $k > 0$), then the doubling period of f is the solution p of the equation $e^{kp} = 2$; that is, $p = (\ln 2)/k$.

Similarly, if $f(x) = ae^{kx}$ is an exponential decay function (so $k < 0$), then the halving period of f is the solution p of the equation $e^{kp} = \frac{1}{2}$; that is, $p = (\ln \frac{1}{2})/k = -(\ln 2)/k$.

Inequalities

An **inequality** consists of two expressions, with one of the four inequality signs between them ($<$, $\leq$, $>$ or $\geq$).

The **solutions** of an inequality are the values of its variables that make the inequality true. These values **satisfy** the inequality.

The **solution set** of an inequality is the set formed by its solutions.

To rearrange an inequality

Carry out any of the following operations on an inequality to obtain an equivalent inequality.

- Rearrange the expressions on one or both sides.
- Swap the sides, *provided you reverse the inequality sign.*
- Do any of the following things to both sides:
  - add or subtract something
  - multiply or divide by something that is positive
  - multiply or divide by something that is negative, *provided you reverse the inequality sign.*

To solve an inequality using a table of signs

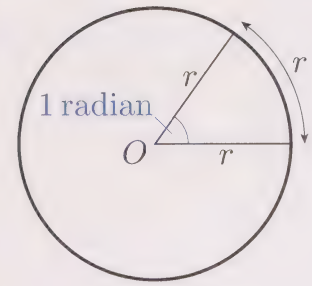
1. Rearrange the inequality to obtain an algebraic expression on one side and only 0 on the other side.
2. If the expression contains any algebraic fractions, then rearrange the whole expression into a single algebraic fraction.
3. Factorise the expression as far as possible.
4. Construct a table of signs:
   (a) In the column headings, write (in increasing order) the values of the variable for which the factors of the expression are equal to zero, and also the largest open intervals to the left and right of, and between, these values.
   (b) In the row headings, write each of the factors of the expression, and then the whole expression.
   (c) For each factor, write 0, + or − in each cell in its row, to indicate whether the factor is zero, positive or negative for the indicated values of the variable.
   (d) Use the signs of the factors to find the signs of the whole expression, and enter these in the bottom row. (Where the expression is undefined, enter the symbol $*$ to indicate this.)
5. Use the entries in the bottom row to solve the inequality.

MST124 Unit 4 Trigonometry

Angles

Radians

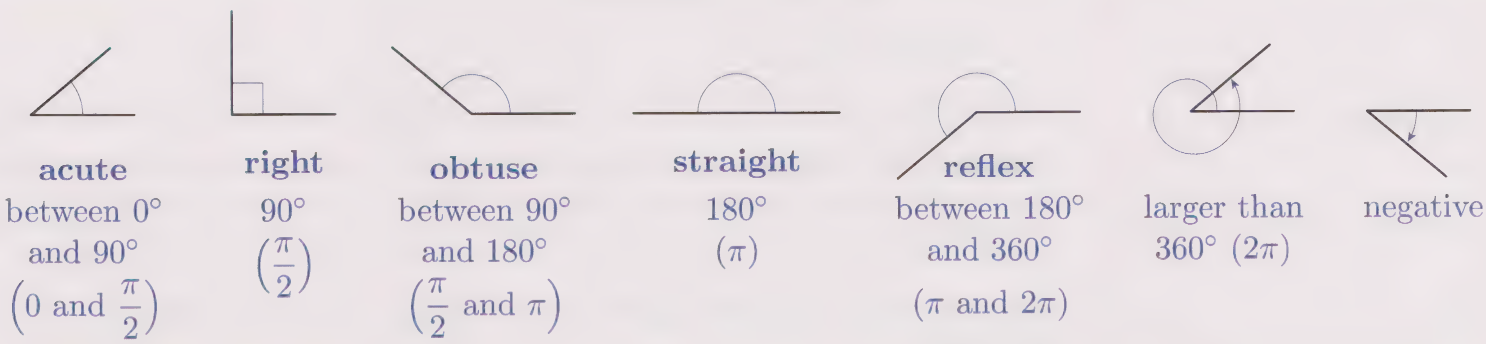
One **radian** is the angle subtended at the centre of a circle by an arc that has the same length as the radius.



So 2π radians $= 360°$.

If the size of an angle is stated with no units, then the angle is measured in radians.

Types of angle



Right-angled triangles

A **right-angled triangle** is a triangle that contains a right angle. Each of the other two angles in the triangle is an acute angle.

The **hypotenuse** of a right-angled triangle is the side opposite the right angle. It is always the longest side.

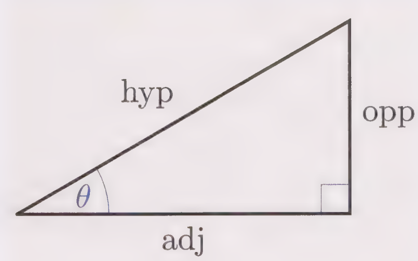
Pythagoras' theorem

For a right-angled triangle, the square of the hypotenuse is equal to the sum of the squares of the other two sides.

Trigonometric ratios for right-angled triangles

The **sine**, **cosine** and **tangent** of the acute angle θ are given by

$$\sin\theta = \frac{\text{opp}}{\text{hyp}}, \quad \cos\theta = \frac{\text{adj}}{\text{hyp}}, \quad \tan\theta = \frac{\text{opp}}{\text{adj}}.$$

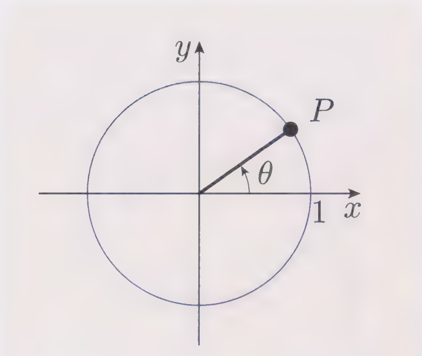


Mnemonic: SOH CAH TOA.

Here hyp, opp and adj denote the hypotenuse, the side opposite θ and the side adjacent to θ, respectively, in a right-angled triangle with an acute angle θ.

For trigonometric ratios for special angles, see page 4.

Trigonometric ratios for angles of any size



The **unit circle** is the circle of radius 1 centred at the origin.

With every angle θ we associate a point P on the unit circle. The point P is obtained by a rotation around the origin through the angle θ, starting from the point on the x-axis with x-coordinate 1. If θ is positive, then the rotation is anticlockwise; if θ is negative, then the rotation is clockwise.

Suppose that θ is any angle and (x, y) are the coordinates of its associated point P on the unit circle. Then

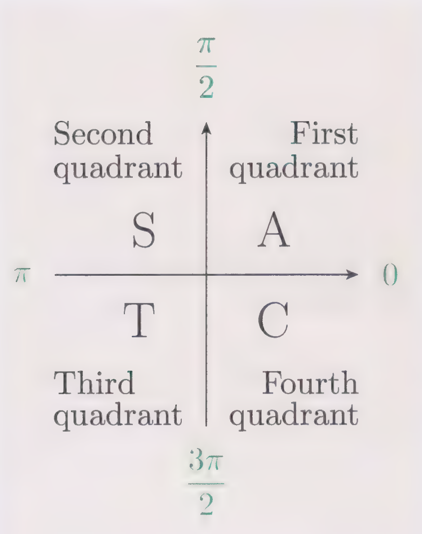
$$\sin\theta = y, \quad \cos\theta = x,$$

and, provided that $x \neq 0$,

$$\tan\theta = \frac{y}{x}.$$

(If $x = 0$, then $\tan\theta$ is undefined.)

The ASTC diagram

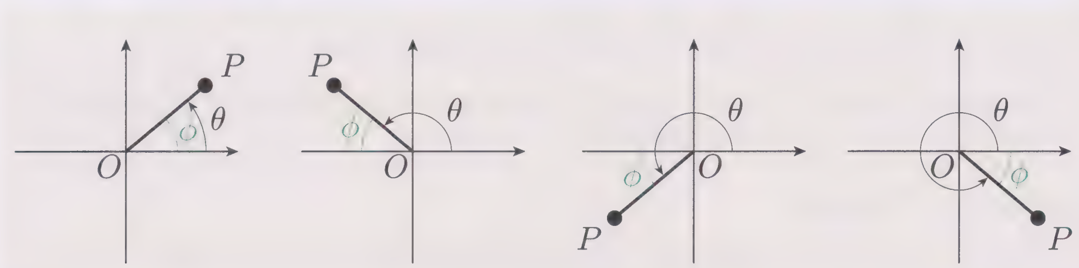


A **quadrant** is one of the four regions separated off by the x- and y-axes.

The **ASTC diagram** indicates which of $\sin\theta$, $\cos\theta$ and $\tan\theta$ are positive when the point P associated with the angle θ lies in each quadrant.
(A stands for all, S stands for sin, T stands for tan, C stands for cos.)

Relationships between trigonometric ratios of angles in different quadrants

Suppose that θ is an angle whose associated point P does not lie on either the x- or y-axis, and ϕ is the acute angle between OP and the x-axis, as in the examples below.



Then

$$\sin\theta = \pm\sin\phi \qquad \cos\theta = \pm\cos\phi \qquad \tan\theta = \pm\tan\phi.$$

The ASTC diagram tells you which sign applies in each case.

(The values $\sin\phi$, $\cos\phi$ and $\tan\phi$ are all positive, because ϕ is acute.)

Two basic trigonometric identities

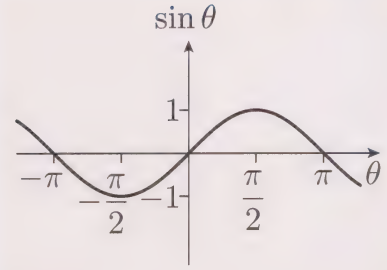
$$\tan\theta = \frac{\sin\theta}{\cos\theta} \qquad \sin^2\theta + \cos^2\theta = 1$$

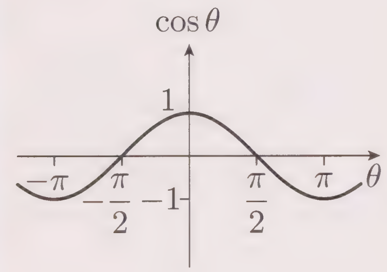
For more trigonometric identities, see page 5.

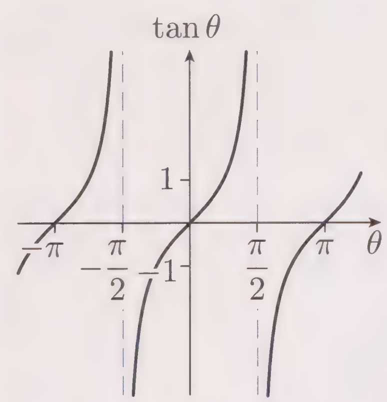
Trigonometric functions

Sine, cosine and tangent functions

The graphs of the sine and cosine functions are **periodic** with **period** 2π (they repeat their shape after every interval of 2π). The graph of the tangent function is periodic with period π.







Inverse sine, cosine and tangent functions

The **inverse sine function** $\sin^{-1}$ has domain $[-1, 1]$ and rule

$$\sin^{-1} x = y,$$

where y is the number in the interval $[-\pi/2, \pi/2]$ such that $\sin y = x$.

The **inverse cosine function** $\cos^{-1}$ has domain $[-1, 1]$ and rule

$$\cos^{-1} x = y,$$

where y is the number in the interval $[0, \pi]$ such that $\cos y = x$.

The **inverse tangent function** $\tan^{-1}$ has domain $\mathbb{R}$ and rule

$$\tan^{-1} x = y,$$

where y is the number in the interval $(-\pi/2, \pi/2)$ such that $\tan y = x$.

The inverse sine, inverse cosine and inverse tangent functions are also called the **arcsine**, **arccosine** and **arctangent** functions.

For the graphs of $\sin^{-1}$, $\cos^{-1}$ and $\tan^{-1}$, see page 10.

Identities for inverse trigonometric functions

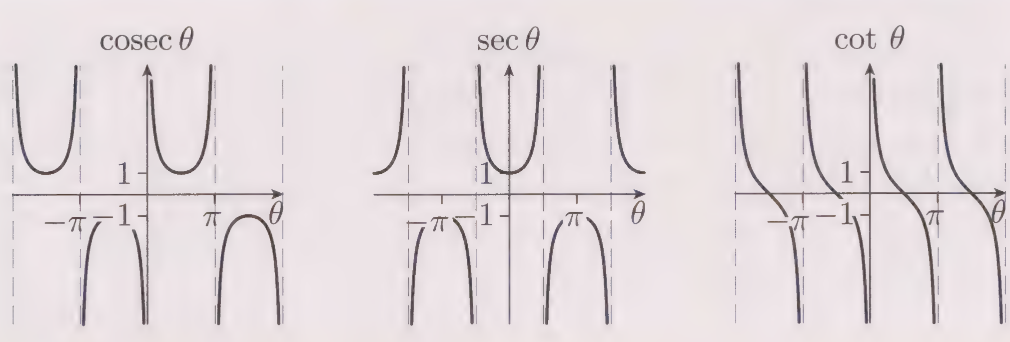
$$\sin^{-1}(-x) = -\sin^{-1} x$$
$$\cos^{-1}(-x) = \pi - \cos^{-1} x$$
$$\tan^{-1}(-x) = -\tan^{-1} x$$

Cosecant, secant and cotangent functions

The **cosecant**, **secant** and **cotangent** functions are given by

$$\operatorname{cosec}\theta = \frac{1}{\sin\theta} \qquad \sec\theta = \frac{1}{\cos\theta} \qquad \cot\theta = \frac{\cos\theta}{\sin\theta}.$$

(So $\cot\theta = 1/(\tan\theta)$, if θ is not an integer multiple of $\pi/2$.)



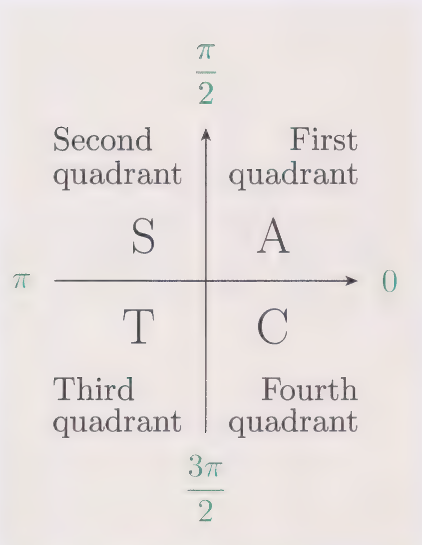
Trigonometric equations

A **trigonometric equation** is an equation that contains a trigonometric function of an unknown.

A trigonometric equation of the form

$$\sin\theta = c, \quad \cos\theta = c \quad \text{or} \quad \tan\theta = c,$$

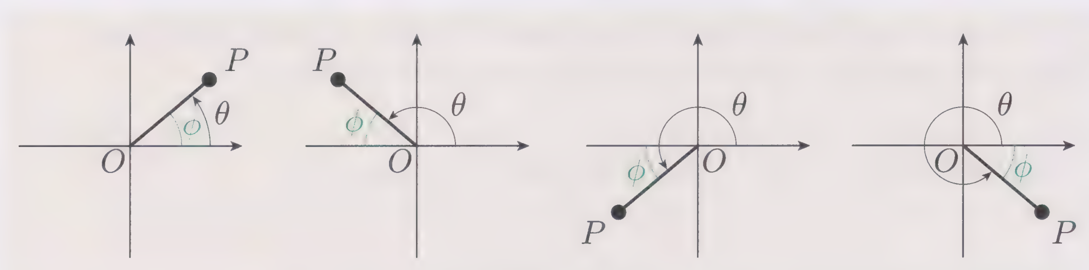
where c is a constant and θ is the unknown, has infinitely many solutions.



To solve an equation of the form $\sin\theta = c$, $\cos\theta = c$ or $\tan\theta = c$ by using the ASTC diagram

(This method does not apply if $c = 0$, or if the equation is $\sin\theta = c$ or $\cos\theta = c$ where $c = \pm 1$.)

1. Use the ASTC diagram to find the quadrants of the solutions (there will be two such quadrants).
2. For each of these quadrants, draw a sketch showing the line OP in that quadrant. On each sketch, mark the associated angle θ that lies in the interval $[0, 2\pi]$, and the acute angle ϕ between OP and the x-axis. (You can use the interval $[-\pi, \pi]$ instead of $[0, 2\pi]$, but you should use the same interval for each sketch.)



3. Find ϕ by applying the appropriate inverse trigonometric function to the equation $\sin\phi = |c|$, $\cos\phi = |c|$ or $\tan\phi = |c|$, as appropriate.
4. Use your sketches to find two values of θ in the interval $[0, 2\pi]$ (or in the interval $[-\pi, \pi]$) that are solutions of the equation.
5. If required, add multiples of 2π to obtain further solutions, or solutions in a different interval.

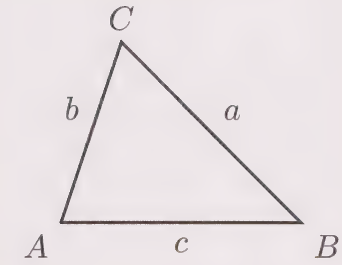
To solve an equation of the form $\sin\theta = c$, $\cos\theta = c$ or $\tan\theta = c$ by using a sketch graph

1. Sketch the graph of the relevant trigonometric function on the interval $[-\pi, \pi]$.
2. Find one solution of the equation by using the appropriate inverse trigonometric function, and mark it on your sketch.
3. Use the symmetry of the graph to find any other solutions in the interval $[-\pi, \pi]$ (usually there is one further such solution).
4. If required, add multiples of 2π to obtain further solutions, or solutions in a different interval.

Trigonometric rules

Sine rule

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}$$



When you use the sine rule to find an unknown angle, you usually obtain two possible angles, one acute and one obtuse. To determine which angle is correct, use additional information about the triangle.

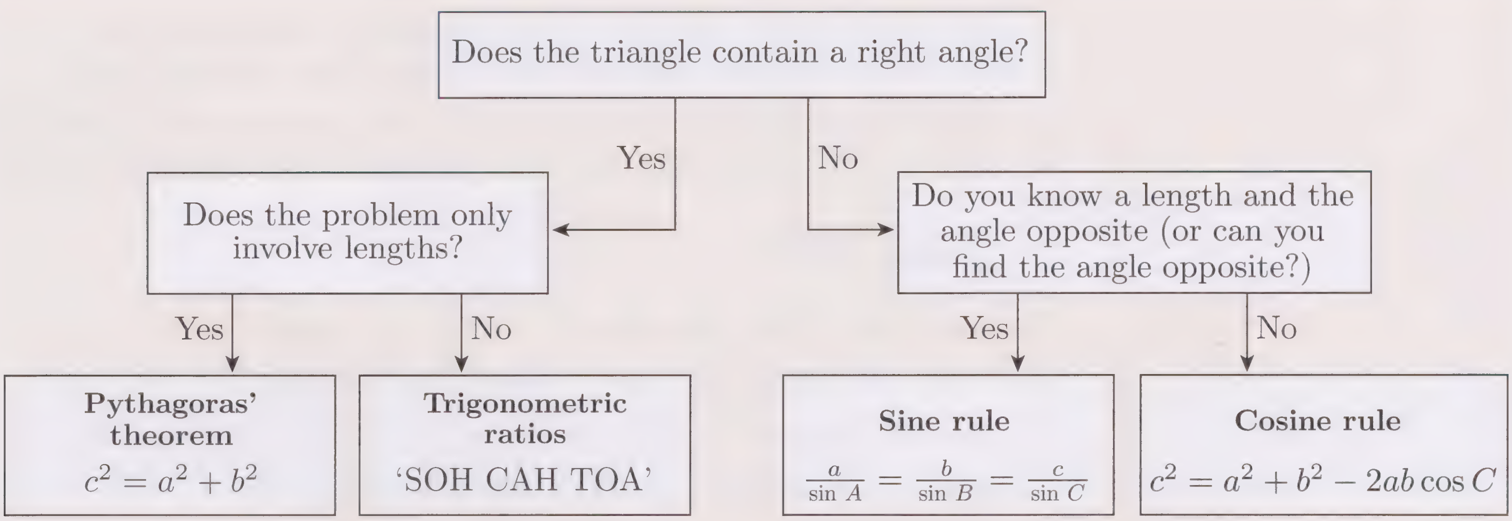
- If the obtuse angle leads to a total angle sum of more than 180°, then it is incorrect.
- Smaller angles are opposite shorter sides.

You may not have enough information to determine which angle is correct.

Cosine rule

$$a^2 = b^2 + c^2 - 2bc\cos A$$
$$b^2 = c^2 + a^2 - 2ca\cos B$$
$$c^2 = a^2 + b^2 - 2ab\cos C$$

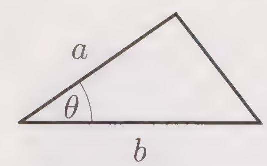
Choosing a method for finding unknown sides and angles in triangles



Area of a triangle

For a triangle with an angle θ between sides of lengths a and b,

$$\text{area} = \tfrac{1}{2}ab\sin\theta.$$

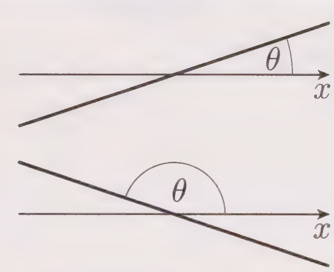


Angle of inclination of a line

The **angle of inclination** of a straight line is the angle that it makes with the x-axis, measured anticlockwise from the positive direction of the x-axis, when the line is drawn on axes *with equal scales.*

For any non-vertical straight line with angle of inclination θ,

$$\text{gradient} = \tan\theta.$$



MST124 Unit 5 Coordinate geometry and vectors

Coordinate geometry

Distance formula (for two or three dimensions)

The distance between the points (x_1, y_1) and (x_2, y_2) is

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

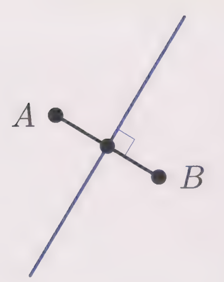
The distance between the points (x_1, y_1, z_1) and (x_2, y_2, z_2) is

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Midpoint formula (for two dimensions)

The midpoint of the line segment joining the points (x_1, y_1) and (x_2, y_2) is

$$\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right).$$

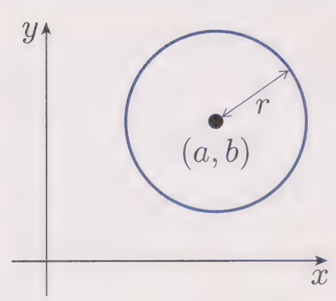


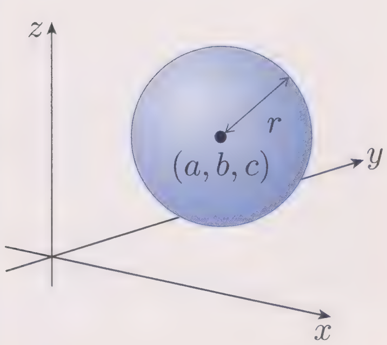
Perpendicular bisectors (for two dimensions)

The **perpendicular bisector** of a line segment is the line that is perpendicular to the line segment and divides it into two equal parts.

If A and B are points in the plane that do not lie on the same horizontal or vertical line, then the gradient of the perpendicular bisector of AB is

$$\frac{-1}{\text{gradient of } AB}.$$





Standard form of the equation of a circle or a sphere

The circle with centre (a, b) and radius r has equation

$$(x - a)^2 + (y - b)^2 = r^2.$$

The sphere with centre (a, b, c) and radius r has equation

$$(x - a)^2 + (y - b)^2 + (z - c)^2 = r^2.$$

To find the equation of the circle passing through three points

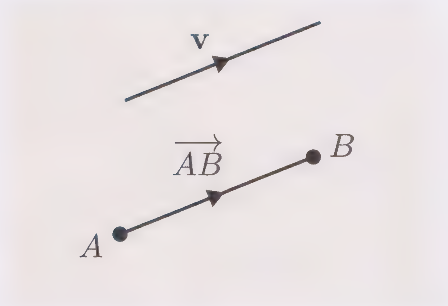
1. Find the equation of the perpendicular bisector of the line segment that joins any pair of the three points.
2. Find the equation of the perpendicular bisector of the line segment that joins a different pair of the three points.
3. Find the point of intersection of these two lines. This is the centre of the circle.
4. Find the radius of the circle, which is the distance from the centre to any of the three points.

Vectors

A **vector** is a quantity that has both a size (usually called **magnitude**) and a direction. The following quantities are vectors:

- **Displacement**, the position of one point relative to another.
- **Velocity**, the speed of an object together with its direction.

A **displacement vector** is a vector that represents displacement.



Scalars

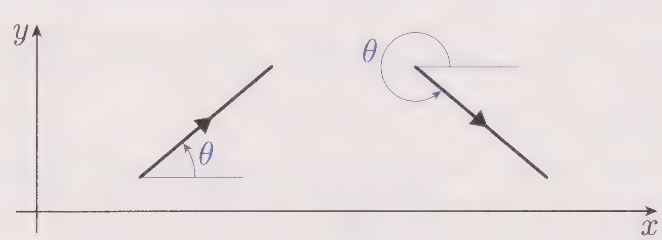
A **scalar** is a quantity that has size but no direction. The following quantities are scalars:

- Distance, the magnitude of displacement.
- Speed, the magnitude of velocity.
- Time, temperature and volume.

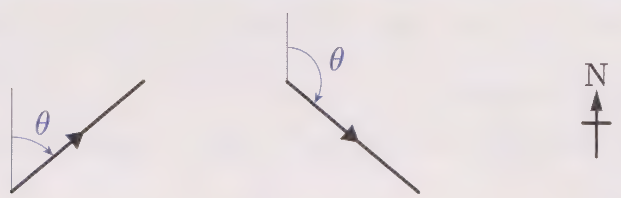
Direction of a vector

The **direction** of a vector is usually specified in one of the following two ways.

- As an angle measured anticlockwise from the positive direction of the x-axis to the direction of the vector.



- As a **bearing**; that is, the angle in degrees between 0° and 360° measured clockwise from north to the direction of the vector.



Heading and course of a ship or aircraft

For a moving ship or aircraft:

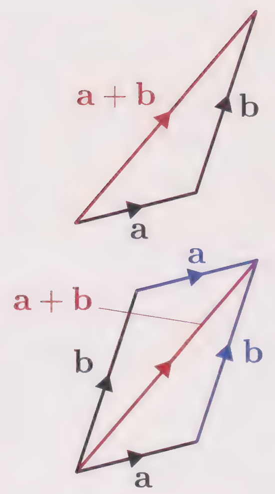
- Its **heading** is the direction in which it is pointing, given as a bearing.
- Its **course** is the direction in which it is actually moving.

These directions may be different, due to the effect of a current or wind.

The actual velocity of a ship or aircraft is the resultant of the velocity that it would have if the water or air were still, and the velocity of the current or wind.

Vector algebra

Two vectors are **equal** if they have the same magnitude and the same direction.



The zero vector

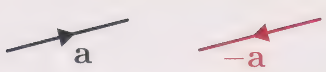
The **zero vector**, denoted by $\mathbf{0}$ (bold zero), is the vector whose magnitude is zero. It has no direction.

Triangle law for vector addition

To find the **sum** (**resultant**) of two vectors $\mathbf{a}$ and $\mathbf{b}$, place the tail of $\mathbf{b}$ at the tip of $\mathbf{a}$. Then $\mathbf{a}+\mathbf{b}$ is the vector from the tail of $\mathbf{a}$ to the tip of $\mathbf{b}$.

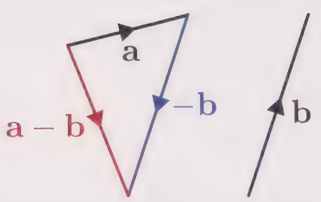
Parallelogram law for vector addition

To find the sum of two vectors $\mathbf{a}$ and $\mathbf{b}$, place their tails together, and complete the resulting figure to form a parallelogram. Then $\mathbf{a}+\mathbf{b}$ is the vector formed by the diagonal of the parallelogram, starting from the point where the tails of $\mathbf{a}$ and $\mathbf{b}$ meet.



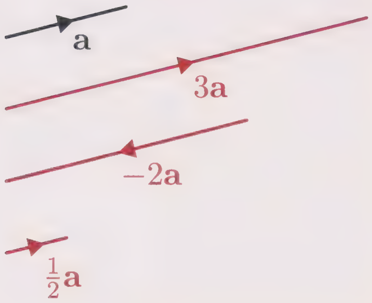
Negative of a vector

The **negative** of a vector $\mathbf{a}$, denoted by $-\mathbf{a}$, is the vector with the same magnitude as $\mathbf{a}$, but the opposite direction.



Vector subtraction

To subtract $\mathbf{b}$ from $\mathbf{a}$, add $-\mathbf{b}$ to $\mathbf{a}$. That is, $\mathbf{a}-\mathbf{b}=\mathbf{a}+(-\mathbf{b})$.



Scalar multiple of a vector

Suppose that $\mathbf{a}$ is a vector. Then, for any non-zero real number m, the **scalar multiple** $m\mathbf{a}$ of $\mathbf{a}$ is the vector

- whose magnitude is $|m|$ times the magnitude of $\mathbf{a}$
- that has the same direction as $\mathbf{a}$ if m is positive, and the opposite direction if m is negative.

Also, $0\mathbf{a}=\mathbf{0}$. (That is, zero times the vector $\mathbf{a}$ is the zero vector.)

Properties of vector algebra

These properties hold for all vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, and all scalars m and n.

1. $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$
2. $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$
3. $\mathbf{a}+\mathbf{0}=\mathbf{a}$
4. $\mathbf{a}+(-\mathbf{a})=\mathbf{0}$
5. $m(\mathbf{a}+\mathbf{b})=m\mathbf{a}+m\mathbf{b}$
6. $(m+n)\mathbf{a}=m\mathbf{a}+n\mathbf{a}$
7. $m(n\mathbf{a})=(mn)\mathbf{a}$
8. $1\mathbf{a}=\mathbf{a}$

Representing vectors using component form

A **unit vector** is a vector with magnitude 1.

The **Cartesian unit vectors**, denoted by $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, are the vectors of magnitude 1 in the directions of the x-, y- and z-axes, respectively.

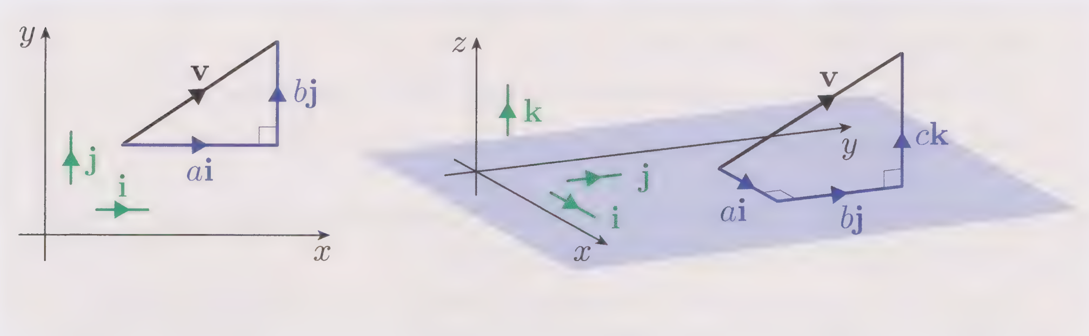
Component form of a vector

The **component form** of a two-dimensional vector $\mathbf{v}$ is the expression $a\mathbf{i} + b\mathbf{j}$, where $\mathbf{v} = a\mathbf{i} + b\mathbf{j}$.

It can also be written as $\begin{pmatrix} a \\ b \end{pmatrix}$.

The **component form** of a three-dimensional vector $\mathbf{v}$ is the expression $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, where $\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$.

It can also be written as $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$.



The **i-component** and **j-component** (or the x-**component** and y-**component**) of a vector $\mathbf{v}$ are the scalars a and b, respectively, in the component form $a\mathbf{i} + b\mathbf{j}$ of $\mathbf{v}$. The components of a three-dimensional vector are referred to in a similar way.

A **column vector** is a vector written as a column, such as $\begin{pmatrix} a \\ b \end{pmatrix}$.

Vector algebra using component form

If $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j}$, and m is a scalar, then

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} \qquad -\mathbf{a} = -a_1\mathbf{i} - a_2\mathbf{j}$$
$$\mathbf{a} - \mathbf{b} = (a_1 - b_1)\mathbf{i} + (a_2 - b_2)\mathbf{j} \qquad m\mathbf{a} = ma_1\mathbf{i} + ma_2\mathbf{j}.$$

In column notation, if $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, and m is a scalar, then

$$\mathbf{a} + \mathbf{b} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix} \qquad -\mathbf{a} = \begin{pmatrix} -a_1 \\ -a_2 \end{pmatrix}$$
$$\mathbf{a} - \mathbf{b} = \begin{pmatrix} a_1 - b_1 \\ a_2 - b_2 \end{pmatrix} \qquad m\mathbf{a} = \begin{pmatrix} ma_1 \\ ma_2 \end{pmatrix}.$$

The algebra of three-dimensional vectors is similar.

Converting vectors from component form to magnitude and direction, and vice versa

To find the magnitude of a two- or three-dimensional vector from its components

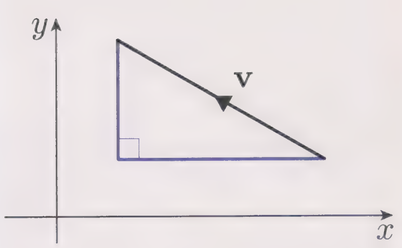
The two-dimensional vector $\mathbf{v} = a\mathbf{i} + b\mathbf{j}$ has magnitude

$$|\mathbf{v}| = \sqrt{a^2 + b^2}.$$

The three-dimensional vector $\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ has magnitude

$$|\mathbf{v}| = \sqrt{a^2 + b^2 + c^2}.$$

To find the magnitude and direction of a two-dimensional vector (not parallel to an axis) from its component form



1. Using the components, sketch a right-angled triangle whose hypotenuse is the vector, and whose shorter sides are parallel to the x- and y-axes.
2. Use Pythagoras' theorem (or, equivalently, the formula above) to find the magnitude of the vector.
3. Use trigonometry to find an acute angle in the triangle.
4. Use this acute angle to work out the direction of the vector.

To find the component form of a two-dimensional vector (not parallel to an axis) from its magnitude and direction

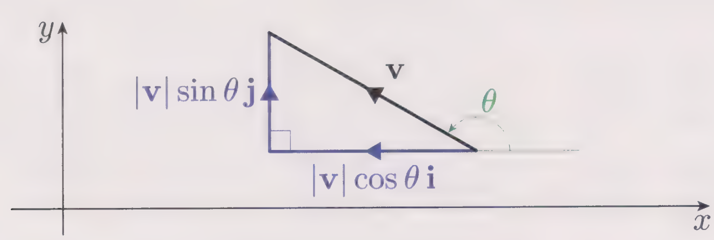
1. Using the magnitude and direction, sketch a right-angled triangle whose hypotenuse is the vector, and whose shorter sides are parallel to the x- and y-axes.
2. Use trigonometry to find the magnitudes of the components.
3. Use the direction of the vector to find the signs of the components.

An alternative method is to use the formula below.

Component form of a two-dimensional vector in terms of its magnitude and the angle that it makes with the positive x-direction

If the two-dimensional vector $\mathbf{v}$ makes the angle θ with the positive x-direction, then

$$\mathbf{v} = |\mathbf{v}|\cos\theta\,\mathbf{i} + |\mathbf{v}|\sin\theta\,\mathbf{j}.$$

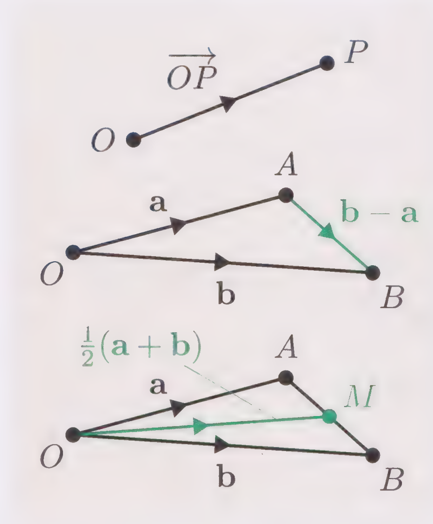


Position vectors

The **position vector** of a point P is the displacement vector $\overrightarrow{OP}$, where O is the origin.

If the points A and B have position vectors $\mathbf{a}$ and $\mathbf{b}$, respectively, then

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a}.$$

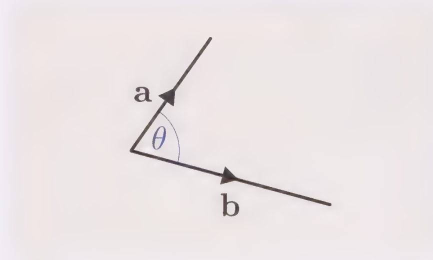


Midpoint formula in terms of position vectors

If the points A and B have position vectors $\mathbf{a}$ and $\mathbf{b}$, respectively, then the midpoint of the line segment AB has position vector $\frac{1}{2}(\mathbf{a}+\mathbf{b})$.

Scalar product

The angle between two non-zero vectors is the angle θ in the range $0 \leq \theta \leq 180°$ between their directions when the vectors are placed tail to tail.



Scalar product of two vectors

The **scalar product** (or **dot product**) of the non-zero vectors $\mathbf{a}$ and $\mathbf{b}$ is

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos\theta,$$

where θ is the angle between $\mathbf{a}$ and $\mathbf{b}$.

If $\mathbf{a}$ or $\mathbf{b}$ is the zero vector, then $\mathbf{a} \cdot \mathbf{b} = 0$.

Properties of the scalar product

These properties hold for all vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, and every scalar m.

1. Suppose that $\mathbf{a}$ and $\mathbf{b}$ are non-zero. If $\mathbf{a}$ and $\mathbf{b}$ are perpendicular, then $\mathbf{a} \cdot \mathbf{b} = 0$, and vice versa.
2. $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$
3. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$
4. $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$
5. $(m\mathbf{a}) \cdot \mathbf{b} = m(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (m\mathbf{b})$

Scalar product in terms of components

If $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j}$, then

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2.$$

If $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, then

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3.$$

To find the angle between two vectors in component form

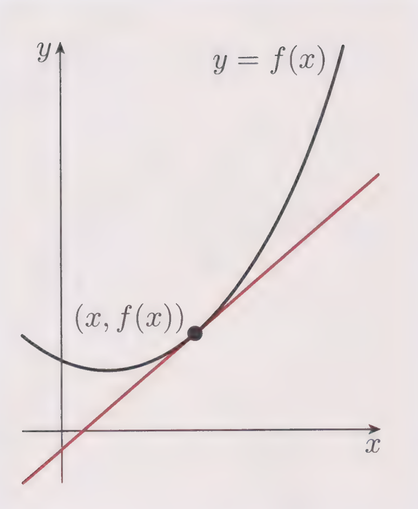
The angle θ between two non-zero vectors $\mathbf{a}$ and $\mathbf{b}$ is given by

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|},$$

where $0 \leq \theta \leq 180°$.

MST124 Unit 6 Differentiation

Derivatives



The **tangent** to a curved graph at a particular point is the straight line that 'just touches' it. The **gradient** of the graph at that point is the gradient of the tangent.

A function f is **differentiable** at a particular value of x if its graph has a gradient at the point $(x, f(x))$.

The **derivative** (or **derived function**) of a function f is the function f' such that

$$f'(x) = \text{gradient of the graph of } f \text{ at the point } (x, f(x)).$$

The domain of f' consists of all the values in the domain of f at which f is differentiable.

The **derivative of** f **at** x is the value $f'(x)$.

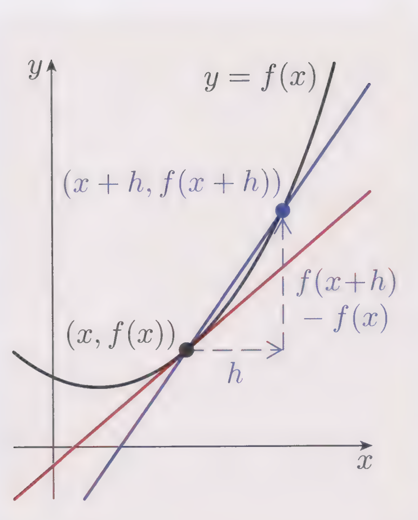
If $y = f(x)$, then $f'(x)$ is the **rate of change** of y with respect to x.

In **Lagrange notation**, the derivative of a function f is denoted by f'.

In **Leibniz notation**, if $y = f(x)$, then $f'(x)$ is denoted by $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ or $\dfrac{\mathrm{d}}{\mathrm{d}x}(f(x))$.

The quantity $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ is called the **derivative of** y **with respect to** x.

Differentiation from first principles



The derivative f' of a function f is given by the equation

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

where the notation '$\lim\limits_{h \to 0}$' means 'the **limit** as h tends to zero of' (the value approached as h approaches zero).

The **difference quotient for** f **at** x is the fraction in the equation above. It is the gradient of the line through the points $(x, f(x))$ and $(x + h, f(x + h))$.

Power functions

A **power function** is a function of the form $f(x) = x^n$, where n is a real number. The power function $f(x) = x^n$ has derivative

$$f'(x) = nx^{n-1}.$$

Constant multiple rule and sum rule for derivatives

If $k(x) = af(x)$, where a is a constant, then

$$k'(x) = af'(x).$$

If $k(x) = f(x) + g(x)$, then

$$k'(x) = f'(x) + g'(x).$$

Increasing and decreasing parts of graphs

Increasing/decreasing criterion

If $f'(x)$ is positive for all x in an interval I, then f is increasing on I.

If $f'(x)$ is negative for all x in an interval I, then f is decreasing on I.

Stationary points

A **stationary point** of a function f is a value of x at which $f'(x) = 0$, or the corresponding point on the graph of f.

- A **local maximum** of f is a point where f takes a value larger than at any other point nearby.
- A **local minimum** of f is a point where f takes a value smaller than at any other point nearby.
- A **horizontal point of inflection** is a stationary point such that the graph is increasing on both sides, or decreasing on both sides.

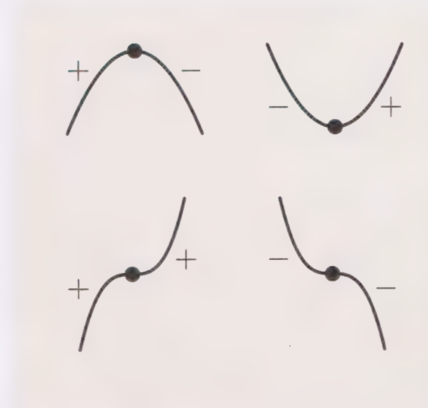
A **turning point** is a local maximum or local minimum.

First derivative test (for determining the nature of a stationary point)

If there are open intervals immediately to the left and right of a stationary point of a function f such that

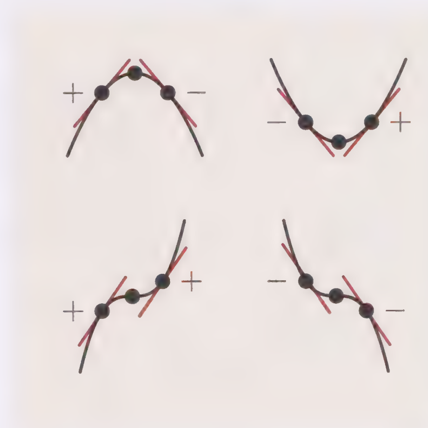
- $f'(x)$ is positive on the left interval and negative on the right interval, then the stationary point is a local maximum
- $f'(x)$ is negative on the left interval and positive on the right interval, then the stationary point is a local minimum
- $f'(x)$ is positive on both intervals or negative on both intervals, then the stationary point is a horizontal point of inflection.

You can apply the first derivative test directly (it can be helpful to construct a table of signs for $f'(x)$), or by choosing sample points.



To apply the first derivative test by choosing sample points

1. Choose two points (that is, two x-values) fairly close to the stationary point, one on each side.
2. Check that the function is differentiable at all points between the chosen points and the stationary point, and that there are no other stationary points between the chosen points and the stationary point.
3. Find the value of the derivative of the function at the two chosen points.
   - If the derivative is positive at the left chosen point and negative at the right chosen point, then the stationary point is a local maximum.
   - If the derivative is negative at the left chosen point and positive at the right chosen point, then the stationary point is a local minimum.
   - If the derivative is positive at both chosen points or negative at both chosen points, then the stationary point is a horizontal point of inflection.



Function continuous on an interval

A function is **continuous** on an interval if its graph has no 'breaks' (**discontinuities**) in the interval.

To find the greatest or least value of a function on an interval of the form $[a, b]$

(This strategy is valid when the function is continuous on the interval, and differentiable at all values in the interval except possibly the endpoints.)

1. Find the stationary points of the function.
2. Find the values of the function at any stationary points inside the interval, and at the endpoints of the interval.
3. Find the greatest or least of the function values found.

Higher derivatives

The **second derivative** (or **second derived function**) of a function f is the function obtained by differentiating f twice. The **third derivative**, the **fourth derivative**, and so on, of f are defined in a similar way.

A function f is n-**times differentiable** at a value x if its nth derivative is defined at x.

Every polynomial function (with domain $\mathbb{R}$) is differentiable infinitely many times at every value of x.

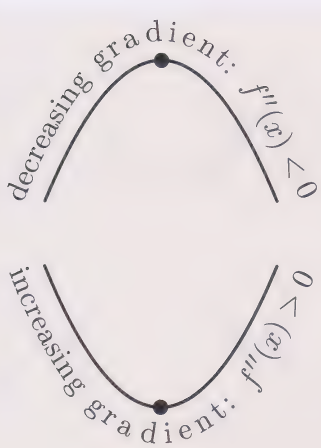
(In fact, as stated in Unit 11, all polynomial, rational, trigonometric, exponential and logarithmic functions, and all constant multiples, sums, differences, products, quotients and composites of these, are differentiable infinitely many times at every value of x in their domains.)

Concave up and concave down

A graph is **concave up** on an interval if the tangents to the graph on that interval lie below the graph. *Mnemonic*: Concave up, like a cup.

A graph is **concave down** on an interval if the tangents to the graph on that interval lie above the graph. *Mnemonic*: Concave down, like a frown.

A **point of inflection** is a point where a graph changes from concave up to concave down or vice versa.



Concave up/concave down criterion

If $f''(x)$ is positive for all x in an interval I, then f is concave up on I.

If $f''(x)$ is negative for all x in an interval I, then f is concave down on I.

Second derivative test (for determining the nature of a stationary point)

If, at a stationary point of a function, the value of the second derivative of the function is

- negative, then the stationary point is a local maximum ☹
- positive, then the stationary point is a local minimum. ☺

Displacement, velocity and acceleration

Suppose that an object is moving along a straight line. If t is the time that has elapsed since some chosen point in time, and s, v and a are the displacement, velocity and acceleration of the object, respectively, then

$$v = \frac{\mathrm{d}s}{\mathrm{d}t}, \quad a = \frac{\mathrm{d}v}{\mathrm{d}t} \quad \text{and} \quad a = \frac{\mathrm{d}^2 s}{\mathrm{d}t^2}.$$

(Time, displacement, velocity and acceleration can be measured in any suitable units, as long as they are consistent.)

Note that the vector quantities displacement, velocity and acceleration are one-dimensional here, so they are represented as scalars, with a plus or minus sign to indicate the direction.

Total cost and marginal cost

The **total cost** of making a particular quantity of a product includes costs that are the same no matter how much of the product is made, and costs that depend on how much of the product is made.

The **unit cost** (or **average cost**) of making a particular quantity of a product is the cost per unit of making that quantity of the product. It is the total cost of making that quantity of the product divided by the number of units of the product made.

The **marginal cost** of making a product is the cost per unit of making more of the product, when a particular quantity of the product is already being made.

The total cost, unit cost and marginal cost are all functions of the quantity of the product made.

The marginal cost is the derivative of the total cost with respect to the quantity of the product made.

MST124 Unit 7 Differentiation methods and integration

For standard derivatives and standard integrals, see page 8.

Differentiation rules

Product rule

Lagrange notation If $k(x) = f(x)g(x)$, then

$$k'(x) = f(x)g'(x) + g(x)f'(x).$$

Leibniz notation If $y = uv$, where u and v are functions of x, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = u\frac{\mathrm{d}v}{\mathrm{d}x} + v\frac{\mathrm{d}u}{\mathrm{d}x}.$$

Informally

$$\begin{pmatrix}\text{derivative}\\ \text{of product}\end{pmatrix} = (\text{first}) \times \begin{pmatrix}\text{derivative}\\ \text{of second}\end{pmatrix} + (\text{second}) \times \begin{pmatrix}\text{derivative}\\ \text{of first}\end{pmatrix}.$$

Quotient rule

Lagrange notation If $k(x) = f(x)/g(x)$, then

$$k'(x) = \frac{g(x)f'(x) - f(x)g'(x)}{(g(x))^2}.$$

Leibniz notation If $y = u/v$, where u and v are functions of x, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2}.$$

Informally

$$\begin{pmatrix}\text{derivative}\\ \text{of quotient}\end{pmatrix} = \frac{(\text{bottom}) \times \begin{pmatrix}\text{derivative}\\ \text{of top}\end{pmatrix} - (\text{top}) \times \begin{pmatrix}\text{derivative}\\ \text{of bottom}\end{pmatrix}}{(\text{bottom})^2}.$$

Chain rule

Lagrange notation If $k(x) = g(f(x))$, then

$$k'(x) = g'(f(x))f'(x).$$

Leibniz notation If y is a function of u, where u is a function of x, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}.$$

Derivative of a function of a linear expression

If $k(x) = f(ax + b)$, where a and b are constants, then

$$k'(x) = af'(ax + b).$$

In particular, if $k(x) = f(ax)$, where a is a constant, then

$$k'(x) = af'(ax).$$

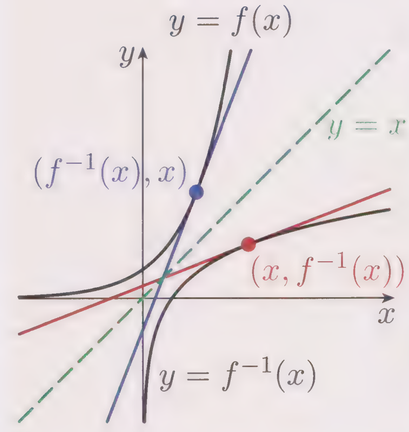
Inverse function rule

Lagrange notation If f is a function with inverse function f^{-1}, then

$$\left(f^{-1}\right)'(x) = \frac{1}{f'(f^{-1}(x))}.$$

Leibniz notation If y is an invertible function of x, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\mathrm{d}x/\mathrm{d}y}.$$



Choosing a method for differentiating a function

1. Is it a standard function (is its derivative given on page 8)?
2. Can you use the constant multiple rule and/or sum rule?
3. Can you rewrite it to make it easier to differentiate?
   (Multiplying out brackets may help.)
4. Is it of the form $f(ax)$ or $f(ax + b)$, where a and b are constants?
   If so, use the rule for differentiating a function of a linear expression.
5. Can you use the product rule?
   Is it of the form $f(x) =$ something $\times$ something?
6. Can you use the quotient rule?
   Is it of the form $f(x) =$ something/something?
7. Can you use the chain rule?
   Is it of the form $f(x) =$ a function of something?

When you use a differentiation rule, you usually have to find the derivatives of simpler functions. Apply the checklist above to each of these simpler functions in turn.

Optimisation problems

An **optimisation problem** involves identifying the best possible option from a choice of suitable possibilities. A **maximisation** or **minimisation problem** involves identifying the circumstances under which the maximum or minimum value, respectively, of a quantity is obtained.

To solve an optimisation problem

1. Identify the quantity that you can change, and represent it by a variable, noting the possible values that it can take. Identify the quantity to be maximised or minimised, and represent it by a variable. These variables are the independent and dependent variables, respectively.
2. Find a formula for the dependent variable in terms of the independent variable.
3. Use the techniques of differential calculus to find the value of the independent variable that gives the maximum/minimum value of the dependent variable. (The strategy for finding the greatest or least value of a function on an interval of the form $[a, b]$, on page 50, is often useful.)
4. Interpret your answer in the context of the problem.

Antiderivatives and indefinite integrals

A **continuous function** is a function whose graph has no 'breaks' (discontinuities) over its whole domain.

An **antiderivative** of a function f is any specific function whose derivative is f.

The **indefinite integral** of a continuous function f is the general function obtained by adding an **arbitrary constant** c (the **constant of integration**) to the formula for an antiderivative of f. It describes the complete family of antiderivatives of f.

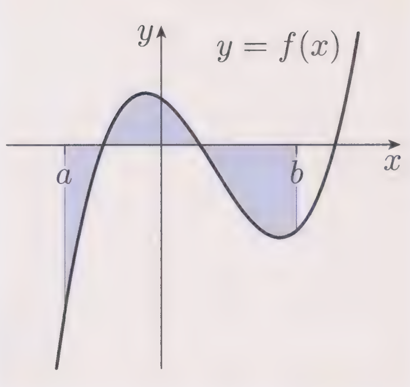
An **indefinite integral** of a function f that is not continuous is a general function obtained by adding an arbitrary constant to the formula for an antiderivative of f. (It can be used to provide the indefinite integral of any continuous function that has the same rule as f.)

Integration (or **antidifferentiation**) is the process of finding an antiderivative of a function. It is the reverse of differentiation.

MST124 Unit 8 Integration methods

Signed areas and definite integrals

Signed areas



The **signed area** of a region that lies entirely above or entirely below the x-axis is its area with a plus or minus sign according to whether it lies above or below the x-axis, respectively.

The **signed area** of a collection of such regions is the sum of the signed areas of the individual regions.

The **signed area between** the graph of a continuous function f and the x-axis **from** $x = a$ **to** $x = b$, where a and b are numbers in its domain, is

- the signed area of the collection of regions described, if $a \leq b$;
- the negative of this signed area, if $b \leq a$.

Definite integrals

The **definite integral** of a continuous function f from a to b, where a and b are numbers in its domain, is the signed area between the graph of f and the x-axis from $x = a$ to $x = b$. It is denoted by

$$\int_a^b f(x)\,\mathrm{d}x.$$

- The numbers a and b are the **lower** and **upper limits of integration**, respectively.
- The expression $f(x)$ is the **integrand**.
- The variable x is a **dummy variable**: you can change its name to any other variable name without affecting the value of the definite integral.

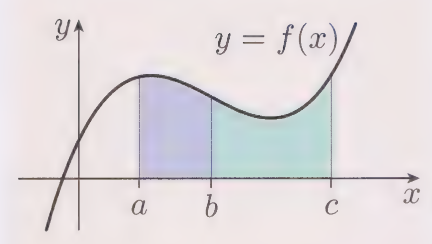
Standard properties of definite integrals

These properties hold for all numbers a, b and c in the domain of a continuous function f.

$$\int_a^a f(x)\,\mathrm{d}x = 0$$

$$\int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x$$

$$\int_a^c f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x$$

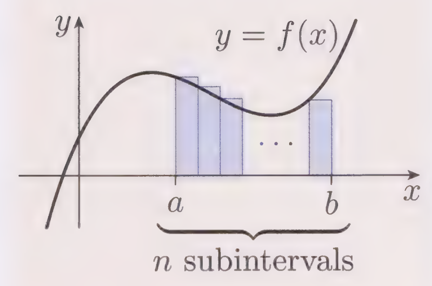


Algebraic definition of a definite integral

The **definite integral** of a continuous function f from $x = a$ to $x = b$, where a and b are numbers in its domain, is given by

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty}\Big(f(a+0w) + f(a+1w) + f(a+2w) + \cdots \\ \cdots + f(a+(n-1)w)\Big)w,$$

where $w = (b-a)/n$.



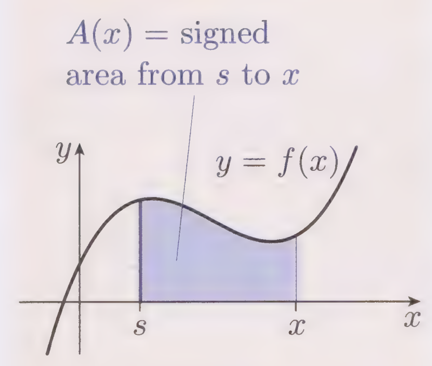
Theorem

Suppose that f is a continuous function, and s is any number in its domain. Let A be the function with the same domain as f and rule

$$A(x) = \int_s^x f(t)\,\mathrm{d}t.$$

Then A is an antiderivative of f.

It follows that: every continuous function has an antiderivative.

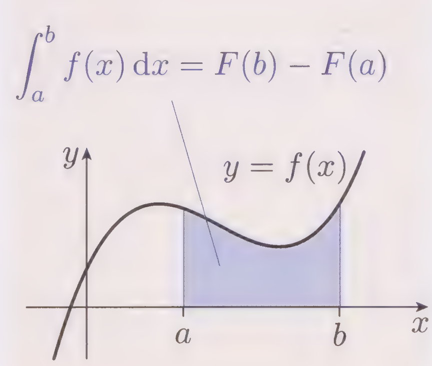


Fundamental theorem of calculus

Suppose that f is a continuous function whose domain contains the numbers a and b, and that F is an antiderivative of f. Then

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a).$$

In **square bracket notation**, $F(b) - F(a)$ is denoted by $\big[F(x)\big]_a^b$.



Notation for indefinite integrals

The indefinite integral of $f(x)$ is denoted by $\displaystyle\int f(x)\,\mathrm{d}x$.

The expression $f(x)$ in this notation is the **integrand**.

Integration rules

Constant multiple rule and sum rule for integrals

For definite integrals

$$\int_a^b k\,f(x)\,\mathrm{d}x = k\int_a^b f(x)\,\mathrm{d}x, \quad \text{where } k \text{ is a constant}$$

$$\int_a^b \big(f(x)+g(x)\big)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x$$

For indefinite integrals

$$\int k\,f(x)\,\mathrm{d}x = k\int f(x)\,\mathrm{d}x, \quad \text{where } k \text{ is a constant}$$

$$\int \big(f(x)+g(x)\big)\,\mathrm{d}x = \int f(x)\,\mathrm{d}x + \int g(x)\,\mathrm{d}x$$

For the square bracket notation

$$\big[kF(x)\big]_a^b = k\big[F(x)\big]_a^b, \quad \text{where } k \text{ is a constant}$$

$$\big[F(x)+G(x)\big]_a^b = \big[F(x)\big]_a^b + \big[G(x)\big]_a^b$$

Integration by substitution

1. Recognise that the integrand is of the form

   $$f(\text{something}) \times \text{the derivative of the something},$$

   where f is a function that you can integrate.
2. Set the something equal to u, and find $\mathrm{d}u/\mathrm{d}x$.
3. Hence write the integral in the form

   $$\int f(u)\,\mathrm{d}u,$$

   by using the fact that $\displaystyle\int f(u)\frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x = \int f(u)\,\mathrm{d}u$.
4. Do the integration.
5. Substitute back for u in terms of x.

For a definite integral, in step 3 also change the limits of integration, which are values of x, to the corresponding values of u. Then omit step 5.

Indefinite integral of a function of a linear expression

Suppose that f is a function with antiderivative F.

If a and b are constants with a non-zero, then

$$\int f(ax+b)\,\mathrm{d}x = \frac{1}{a}F(ax+b)+c.$$

In particular, if a is a non-zero constant, then

$$\int f(ax)\,\mathrm{d}x = \frac{1}{a}F(ax)+c.$$

Integration by parts

Lagrange notation $\displaystyle\int f(x)g(x)\,\mathrm{d}x = f(x)G(x) - \int f'(x)G(x)\,\mathrm{d}x.$

Here G is an antiderivative of g.

Alternative version $\displaystyle\int f(x)g'(x)\,\mathrm{d}x = f(x)g(x) - \int f'(x)g(x)\,\mathrm{d}x$

Leibniz notation $\displaystyle\int u\,\frac{\mathrm{d}v}{\mathrm{d}x}\,\mathrm{d}x = uv - \int v\,\frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x$

Informally

$$\begin{aligned}\begin{pmatrix}\text{integral}\\ \text{of product}\end{pmatrix} &= (\text{first}) \times \begin{pmatrix}\text{antiderivative}\\ \text{of second}\end{pmatrix}\\ &\quad - \text{integral of}\left(\begin{pmatrix}\text{derivative}\\ \text{of first}\end{pmatrix} \times \begin{pmatrix}\text{antiderivative}\\ \text{of second}\end{pmatrix}\right).\end{aligned}$$

For definite integrals $\displaystyle\int_a^b f(x)g(x)\,\mathrm{d}x = \Big[f(x)G(x)\Big]_a^b - \int_a^b f'(x)G(x)\,\mathrm{d}x.$

Here G is an antiderivative of g.

Choosing a method for finding an integral

- Is it a standard integral? Consult the table on page 8.
- Can you rearrange the integral to express it as a sum of constant multiples of simpler integrals?
- Is the integrand of the form $f(ax)$ or $f(ax+b)$, where a and b are constants and f is a function that you can integrate? If so, use the rule for integrating a function of a linear expression, or use integration by substitution, with $u = ax$ or $u = ax + b$ as appropriate.
- Can the integrand be written in the form

  $f(\text{something}) \times \text{the derivative of the something},$

  where f is a function that you can integrate? (To obtain this form you might need to multiply by a constant inside the integral, and divide by the same constant outside.) If so, use integration by substitution. Start by setting the 'something' equal to u.
- Is the integrand of the form $f(x)g(x)$, where f is a function that becomes simpler when differentiated, and g is a function that you can integrate? (For example, $f(x)$ might be x, or x^2, or any polynomial expression in x, or it might be $\ln x$.) If so, try integration by parts.
- If the integrand contains trigonometric functions, can you use trigonometric identities to rewrite it in a form that's easier to integrate?

To integrate an expression of the form $e^{ax}\sin(bx)$ or $e^{ax}\cos(bx)$

Integrate by parts twice, to obtain an equation that expresses the original integral in terms of itself, then rearrange this to find the original integral.

MST124 Unit 9 Matrices

Matrices and matrix operations

A **matrix** is a rectangular array of numbers. An **element** (or **entry**) of a matrix is one of the numbers in it. The element in row i and column j of a matrix $\mathbf{A}$ is denoted by a_{ij}.

A **row** of a matrix is a horizontal line of numbers in it, and a **column** of a matrix is a vertical line of numbers in it. An $m \times n$ matrix, or matrix of **size** $m \times n$, has m rows and n columns.

A **square matrix** has the same number of rows as columns.

A **vector** is a matrix with one column. The **components** of a vector are its elements. An n-**dimensional** vector has n components.

Zero matrix

A **zero matrix** is a matrix each of whose elements is zero.

Negative of a matrix

The **negative** of a matrix $\mathbf{A}$, denoted by $-\mathbf{A}$, is the matrix obtained by changing the sign of each element of $\mathbf{A}$.

Matrix addition and subtraction

If $\mathbf{A}$ and $\mathbf{B}$ are matrices of the same size, then $\mathbf{A}+\mathbf{B}$ is the matrix obtained by adding the corresponding elements of $\mathbf{A}$ and $\mathbf{B}$, and $\mathbf{A}-\mathbf{B}$ is the matrix obtained by subtracting from each element of $\mathbf{A}$ the corresponding element of $\mathbf{B}$. Two matrices of different sizes cannot be added or subtracted.

Scalar multiplication of a matrix

If $\mathbf{A}$ is a matrix and k is a number, then $k\mathbf{A}$ is the matrix obtained by multiplying each element of $\mathbf{A}$ by k.

Properties of matrix addition and scalar multiplication

These properties hold for all matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ of the same size, and all scalars m and n.

1. $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$
2. $(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})$
3. $\mathbf{A}+\mathbf{0}=\mathbf{A}$
4. $\mathbf{A}+(-\mathbf{A})=\mathbf{0}$
5. $m(\mathbf{A}+\mathbf{B})=m\mathbf{A}+m\mathbf{B}$
6. $(m+n)\mathbf{A}=m\mathbf{A}+n\mathbf{A}$
7. $m(n\mathbf{A})=(mn)\mathbf{A}$
8. $1\mathbf{A}=\mathbf{A}$

In properties 3 and 4, $\mathbf{0}$ is the zero matrix of the same size as $\mathbf{A}$.

Matrix multiplication

If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is an $n \times p$ matrix, then $\mathbf{AB}$ is the $m \times p$ matrix in which, for each i and j, the entry in row i and column j is obtained by multiplying each element in the ith row of $\mathbf{A}$ by the corresponding element in the jth column of $\mathbf{B}$, and adding the results.

$$\begin{array}{cccccc} & & \text{column } j & & \text{column } j & \\ \text{row } i & \begin{pmatrix} & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ & & & \end{pmatrix} & \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} & = & \begin{pmatrix} & \vdots & \\ \cdots & \bigcirc & \cdots \\ & \vdots & \end{pmatrix} & \text{row } i \end{array}$$

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

If the number of columns of a matrix $\mathbf{A}$ is not equal to the number of rows of a matrix $\mathbf{B}$, then $\mathbf{AB}$ is not defined.

Properties of matrix multiplication

These properties hold for all matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ for which the products and sums mentioned are defined.

1. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
2. $k(\mathbf{AB}) = (k\mathbf{A})\mathbf{B} = \mathbf{A}(k\mathbf{B})$, for any scalar k
3. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
4. $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$

For matrices $\mathbf{A}$ and $\mathbf{B}$:

- The product $\mathbf{AB}$ may exist but the product $\mathbf{BA}$ may not.
- If both $\mathbf{AB}$ and $\mathbf{BA}$ exist, they are usually not equal.

Matrix powers

The **square** $\mathbf{A}^2$ of a square matrix $\mathbf{A}$ is the matrix product $\mathbf{AA}$.

The nth **power** $\mathbf{A}^n$ of a square matrix $\mathbf{A}$ is is obtained by multiplying together n copies of $\mathbf{A}$. For example, $\mathbf{A}^3 = \mathbf{AAA}$.

Matrix inverses and determinants

Identity matrices

An **identity matrix** is a square matrix $\mathbf{I}$ of the form $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$.

If $\mathbf{I}$ is an identity matrix, then

- for any matrix $\mathbf{A}$ for which the product $\mathbf{AI}$ is defined, $\mathbf{AI} = \mathbf{A}$
- for any matrix $\mathbf{A}$ for which the product $\mathbf{IA}$ is defined, $\mathbf{IA} = \mathbf{A}$.

Inverse of a matrix

A square matrix $\mathbf{A}$ is **invertible** if there is a matrix $\mathbf{B}$ of the same size such that $\mathbf{AB} = \mathbf{I}$ and $\mathbf{BA} = \mathbf{I}$, where $\mathbf{I}$ is an identity matrix. For each square matrix $\mathbf{A}$ there is at most one such matrix $\mathbf{B}$; if there is such a matrix $\mathbf{B}$, then it is called the **inverse** of $\mathbf{A}$ and denoted by $\mathbf{A}^{-1}$.

So, if $\mathbf{A}$ is an invertible matrix, then

$$\mathbf{AA}^{-1} = \mathbf{I} \quad \text{and} \quad \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

A square matrix is **non-invertible** if it has no inverse.

Determinant of a matrix

The **determinant** of a square matrix $\mathbf{A}$ is a number calculated from its elements. It is denoted by $\det \mathbf{A}$.

- If $\det \mathbf{A} \neq 0$, then $\mathbf{A}$ is invertible.
- If $\det \mathbf{A} = 0$, then $\mathbf{A}$ is not invertible.

Determinant and inverse of a 2×2 matrix

If $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then

- $\det \mathbf{A} = ad - bc$
- $\mathbf{A}^{-1} = \dfrac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, provided that $ad - bc \neq 0$.

Systems of linear equations

A **system of linear equations** is a set of linear equations, each in the same set of unknowns, that apply simultaneously.

A **solution** of the system is an assignment of values to the unknowns that makes all the equations true simultaneously.

Matrix form of a system of linear equations

A system of n linear equations in n unknowns can be written as a single matrix equation

$$\mathbf{Ax} = \mathbf{b},$$

where

- $\mathbf{A}$ is an $n \times n$ matrix called the **coefficient matrix**
- $\mathbf{x}$ is an n-dimensional vector whose components are the unknowns
- $\mathbf{b}$ is an n-dimensional vector.

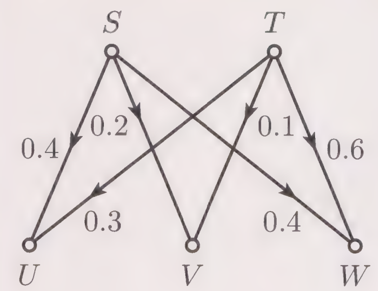
The system $\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned}$ has matrix form $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix}$.

Solutions of a system of linear equations

- If $\det \mathbf{A} \neq 0$, then the system $\mathbf{Ax} = \mathbf{b}$ has a unique solution, given by $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.
- If $\det \mathbf{A} = 0$, then the system $\mathbf{Ax} = \mathbf{b}$ has no solution or infinitely many solutions.

Matrices and networks

A network with n input nodes and m output nodes, in which all flow is directly from input nodes to output nodes, can be represented by an $m \times n$ matrix $\mathbf{A}$ in which, for each i and j, the element a_{ij} is the proportion of the flow from the jth input node that goes to the ith output node. (There is an example in the margin.)



$$\begin{array}{c} \\ U \\ V \\ W \end{array}\begin{array}{c} \begin{array}{cc} S & T \end{array} \\ \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} \end{array} = \mathbf{A}$$

If the n input values are represented by an n-dimensional vector $\mathbf{x}$, and the corresponding m output values are represented by an m-dimensional vector $\mathbf{y}$, then $\mathbf{y} = \mathbf{Ax}$.

Combining networks

If the output nodes of one network are also the input nodes of a second network, then the combined network is equivalent to a simpler network in which flow is directly from the input nodes of the first network to the output nodes of the second network.

This simpler network is represented by the matrix $\mathbf{BA}$, where $\mathbf{A}$ and $\mathbf{B}$ are the matrices representing the first and second networks respectively.

MST124 Unit 10 Sequences and series

Sequences

A **sequence** is a list of numbers, called its **terms**.

A **finite sequence** has a finite number of terms.

An **infinite sequence** has an infinite number of terms.

The notation (a_n) denotes the infinite sequence $a_1, a_2, a_3, \ldots$.
The variable n is the **index variable**.

The first term of a sequence is assumed to have subscript 1, unless otherwise indicated.

A **constant sequence** is a sequence in which every term has the same value.

Closed form for a sequence

A **closed form** for a sequence is a formula that defines the general term a_n as an expression involving the subscript n. To specify a sequence using a closed form, two pieces of information are needed:

- the closed form
- the range of values for the subscript n.

Recurrence system for a sequence

A (**first-order**) **recurrence relation** for a sequence is an equation that defines each term a_n other than the first as an expression involving the previous term a_{n-1}. To specify a sequence using a **recurrence system**, three pieces of information are needed:

- the value of the first term
- the recurrence relation
- the range of values for the subscript n.

Arithmetic sequences

An **arithmetic sequence** is a sequence in which each term (apart from the first) is obtained by adding a fixed number, called the **common difference**, to the previous term.

The arithmetic sequence (x_n) with first term a and common difference d has recurrence system

$$x_1 = a, \qquad x_n = x_{n-1} + d \quad (n = 2, 3, 4, \ldots)$$

and closed form

$$x_n = a + (n-1)d \quad (n = 1, 2, 3, \ldots).$$

An **arithmetic series** is an expression obtained by adding consecutive terms of an arithmetic sequence.

Geometric sequences

A **geometric sequence** is a sequence in which each term (apart from the first) is obtained by multiplying the previous term by a fixed number, called the **common ratio**.

The geometric sequence (x_n) with first term a and common ratio r has recurrence system

$$x_1 = a, \qquad x_n = r x_{n-1} \quad (n = 2, 3, 4, \ldots)$$

and closed form

$$x_n = ar^{n-1} \quad (n = 1, 2, 3, \ldots).$$

A **geometric series** is an expression obtained by adding consecutive terms of a geometric sequence.

Long-term behaviour of infinite sequences

Types of long-term behaviour

A sequence (x_n)

- is **increasing** if $x_{n-1} < x_n$ for each pair of successive terms x_{n-1} and x_n
- is **decreasing** if $x_{n-1} > x_n$ for each pair of successive terms x_{n-1} and x_n
- is **bounded** if all the terms lie within some interval $[-A, A]$, where A is a fixed positive number
- is **unbounded** if there is no fixed value of A, however large, for which all the terms lie within the interval $[-A, A]$
- **converges** to the **limit** L if its terms approach L more and more closely, so that eventually they lie within any interval $[L-h, L+h]$, no matter how small the positive number h is taken to be. We say that x_n **tends to** L **as** n **tends to infinity**, and write

  $$x_n \to L \text{ as } n \to \infty \quad \text{or} \quad \lim_{n\to\infty} x_n = L$$

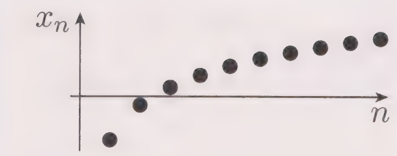
- **tends to infinity** if its terms increase without limit, so that eventually they lie in any interval $[A, \infty)$, no matter how large the positive number A is taken to be. We say that x_n **tends to infinity as** n **tends to infinity**, and write

  $$x_n \to \infty \text{ as } n \to \infty$$

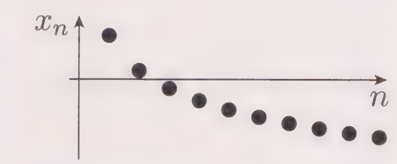
- **tends to minus infinity** if its terms decrease without limit, so that eventually they lie in any interval $(-\infty, -A]$, no matter how large the positive number A is taken to be. We say that x_n **tends to minus infinity as** n **tends to infinity**, and write

  $$x_n \to -\infty \text{ as } n \to \infty.$$

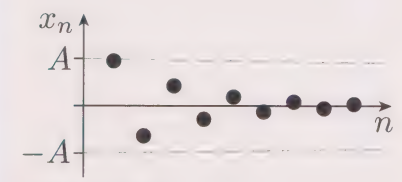
Some examples are shown in the margin.



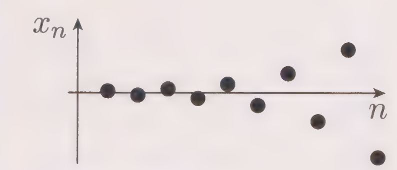
Increasing



Decreasing



Bounded

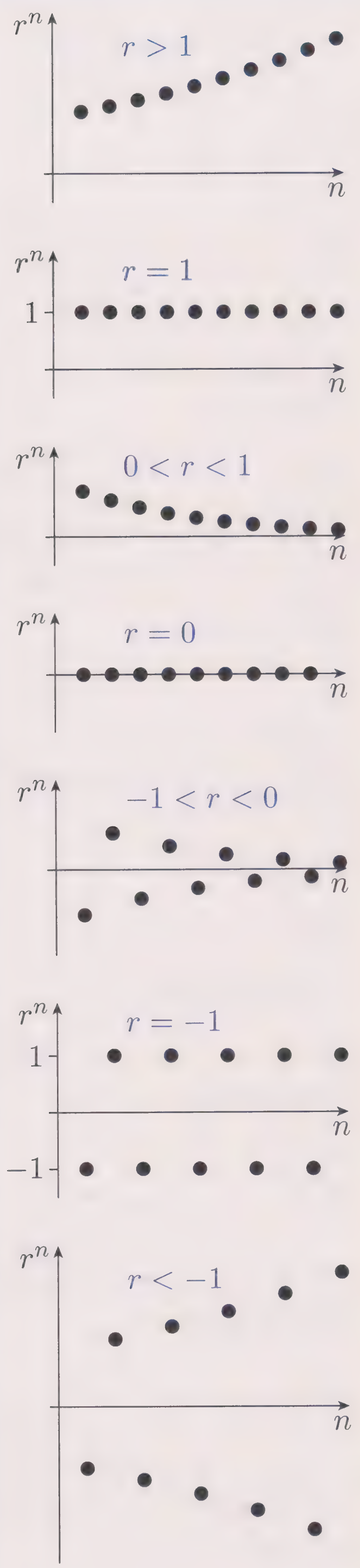


Unbounded

Converging to L

Tending to infinity

Tending to minus infinity



Behaviour of (r^n)

Long-term behaviour of arithmetic sequences

Suppose that (x_n) is an arithmetic sequence with common difference d.

- If $d > 0$, then (x_n) is increasing and $x_n \to \infty$ as $n \to \infty$.
- If $d < 0$, then (x_n) is decreasing and $x_n \to -\infty$ as $n \to \infty$.
- If $d = 0$, then (x_n) is constant.

Long-term behaviour of the sequence (r^n)

Value of r	Behaviour of (r^n)
$r > 1$	Increasing, $r^n \to \infty$ as $n \to \infty$
$r = 1$	Constant: $1, 1, 1, \ldots$
$0 < r < 1$	Decreasing, $r^n \to 0$ as $n \to \infty$
$r = 0$	Constant: $0, 0, 0, \ldots$
$-1 < r < 0$	Alternates in sign, $r^n \to 0$ as $n \to \infty$
$r = -1$	Alternates between -1 and 1
$r < -1$	Alternates in sign, unbounded

Effects of multiplying each term by a constant

Suppose that (x_n) is a sequence and c is a constant.

- If $c \neq 0$ and (x_n) $\left\{\begin{array}{l}\text{is constant}\\ \text{alternates in sign}\\ \text{is bounded}\\ \text{is unbounded}\\ \text{tends to } 0\end{array}\right\}$, then so is/does (cx_n).
- If $c > 0$ and (x_n) $\left\{\begin{array}{l}\text{is increasing}\\ \text{is decreasing}\\ \text{tends to } \infty\\ \text{tends to } -\infty\end{array}\right\}$, then so is/does (cx_n).
- If $c < 0$ and (x_n) $\left\{\begin{array}{l}\text{is increasing}\\ \text{is decreasing}\\ \text{tends to } \infty\\ \text{tends to } -\infty\end{array}\right\}$, then (cx_n) $\left\{\begin{array}{l}\text{is decreasing}\\ \text{is increasing}\\ \text{tends to } -\infty\\ \text{tends to } \infty\end{array}\right\}$.
- If $x_n \to L$ as $n \to \infty$, then $cx_n \to cL$ as $n \to \infty$.

Effects of adding a constant to each term

Suppose that (x_n) is a sequence and a and L are constants.

- If $x_n \to L$ as $n \to \infty$, then $x_n + a \to L + a$ as $n \to \infty$.
- If $x_n \to \infty$ as $n \to \infty$, then $x_n + a \to \infty$ as $n \to \infty$.
- If $x_n \to -\infty$ as $n \to \infty$, then $x_n + a \to -\infty$ as $n \to \infty$.

Series

A **series** is an expression obtained by adding consecutive terms of a sequence.

Sums of series

The **sum** of a finite series is the number obtained by adding up all the terms in the series.

The nth **partial sum** s_n of an infinite series $a_1 + a_2 + a_3 + \cdots$ is the number obtained by adding up the first n terms; $s_n = a_1 + a_2 + \cdots + a_n$.

An infinite series $a_1 + a_2 + a_3 + \cdots$ has **sum** s if its sequence of partial sums (s_n) converges to the limit s.

For formulas for sums of standard finite series, see page 9.

Sigma notation

- The finite sum $x_p + x_{p+1} + \cdots + x_q$ is denoted by $\displaystyle\sum_{n=p}^{q} x_n$.
- The infinite sum $x_p + x_{p+1} + \cdots$ is denoted by $\displaystyle\sum_{n=p}^{\infty} x_n$.

The variable n is the **index variable**. It is a dummy variable.
The numbers p and q are the **lower** and **upper limits**, respectively.

Rules for manipulating finite series in sigma notation

$$\sum_{k=p}^{q} cx_k = c\sum_{k=p}^{q} x_k \quad \text{(where } c \text{ is a constant)}$$

$$\sum_{k=p}^{q}(x_k + y_k) = \sum_{k=p}^{q} x_k + \sum_{k=p}^{q} y_k$$

$$\sum_{k=p}^{q} x_k = \sum_{k=1}^{q} x_k - \sum_{k=1}^{p-1} x_k \quad \text{(where } 1 < p \leq q\text{)}$$

These rules also apply when q is replaced by ∞, provided that each series involved has a sum.

The binomial theorem

For any natural number n,

$$\begin{aligned}(a+b)^n &= {}^nC_0\, a^n + {}^nC_1\, a^{n-1}b + \cdots + {}^nC_{n-1}\, ab^{n-1} + {}^nC_n\, b^n \\ &= \sum_{k=0}^{n} {}^nC_k\, a^{n-k}b^k,\end{aligned}$$

where the **binomial coefficient** ${}^nC_k$ is given by

$${}^nC_k = \frac{n!}{k!\,(n-k)!} \quad \text{for } k = 0, 1, 2, \ldots, n.$$

$$\left(\text{For } 0 < k \leq n, \text{ alternatively } {}^nC_k = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1}.\right)$$

Here $n!$ denotes n **factorial**. If $n \geq 1$ then $n! = 1 \times 2 \times \cdots \times n$; $0! = 1$.

MST124 Unit 11 Taylor polynomials

Taylor polynomials

Let f be a function that is n-times differentiable at a point a. The **Taylor polynomial of degree n about a for f** is

$$p(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

The point a is called the **centre** of the Taylor polynomial.

When $a = 0$, the Taylor polynomial becomes

$$p(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n.$$

A Taylor polynomial of degree n can be denoted by $p_n(x)$.

The function f and the Taylor polynomial p of degree n about a for f have the same value at $x = a$, and the first, second, third, ..., nth derivatives of f have the same values at $x = a$ as the corresponding derivatives of p.

(Note that we use **point** to mean 'number'.)

Taylor series

Let f be a function that is differentiable infinitely many times at a point a. The **Taylor series about a for f** is

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots.$$

The point a is called the **centre** of the Taylor series.

When $a = 0$, the Taylor series becomes

$$f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots.$$

The **Maclaurin series** for a function f is its Taylor series about 0.

Validity of Taylor series

A Taylor series about a point a for a function f is **valid** at a point x if for that value of x the Taylor series has sum $f(x)$.

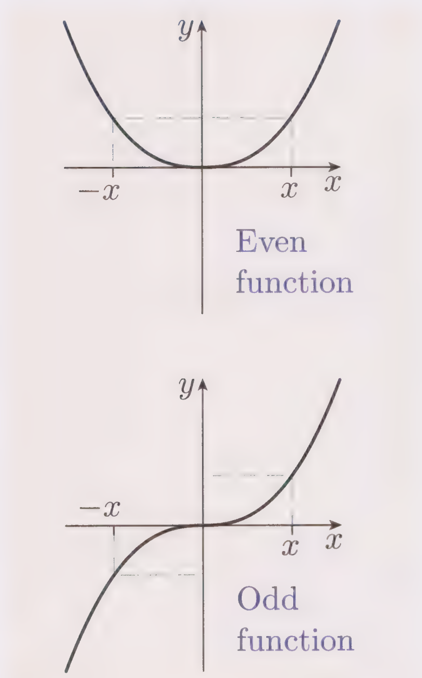
An **interval of validity** for a Taylor series about a point a for a function f is an interval of points for which the Taylor series is valid.

Standard Taylor series

For some standard Taylor series about 0, see page 9.

The **binomial series** is the standard Taylor series about 0 for $(1+x)^\alpha$.

Even and odd functions



A function f is **even** if $f(-x) = f(x)$ for all x in the domain of f.

The graph of an even function is unchanged under reflection in the y-axis.

A function f is **odd** if $f(-x) = -f(x)$ for all x in the domain of f.

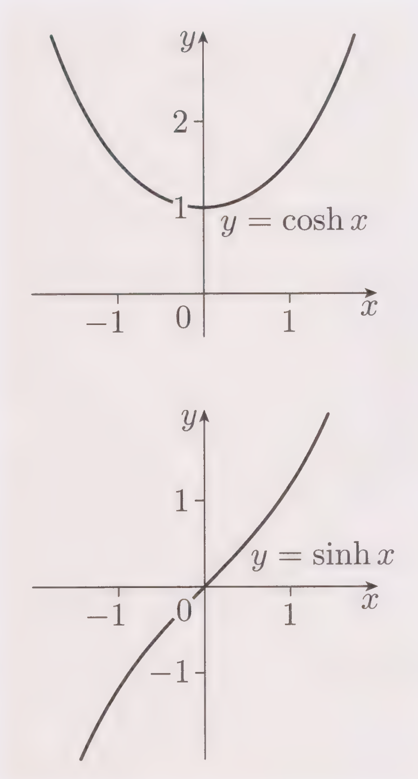
The graph of an odd function is unchanged by a rotation through a half-turn about the origin.

The cosine function is an even function.
The sine function is an odd function.

Taylor polynomials and series about 0 for even and odd functions

- A Taylor polynomial or Taylor series about 0 for an even function contains terms in even powers of x only.
- A Taylor polynomial or Taylor series about 0 for an odd function contains terms in odd powers of x only.

Hyperbolic functions



The **hyperbolic cosine function** is given by

$$\cosh x = \tfrac{1}{2}(e^x + e^{-x}).$$

The **hyperbolic sine function** is given by

$$\sinh x = \tfrac{1}{2}(e^x - e^{-x}).$$

The hyperbolic cosine function is an even function.
The hyperbolic sine function is an odd function.

The Taylor series about 0 for these functions are:

$$\cosh x = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \frac{1}{6!}x^6 + \dots, \quad \text{for } x \in \mathbb{R}$$

$$\sinh x = x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \frac{1}{7!}x^7 + \dots, \quad \text{for } x \in \mathbb{R}.$$

Using Taylor polynomials for approximation

To find an approximation for $f(x)$, for a particular function f at a particular point x

1. Find a Taylor series for f about a suitable point a close to x, and check that it is valid at x.
2. Truncate the Taylor series to obtain a Taylor polynomial, and evaluate it at x.

To find an approximation for $f(x)$, for a particular function f at a particular point x, to m decimal places (rule of thumb)

Use the method above to calculate approximations by using Taylor polynomials of degrees 1, 2, 3, and so on, until two successive different approximations agree to $m + 2$ decimal places.

Uniqueness of Taylor series

Let f be a function. If you can by any means find a series

$$c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \cdots$$

that is equal to $f(x)$ for all x in some open interval containing a, then this series is the Taylor series about a for f, and hence it is the *only* series of this form that is equal to $f(x)$ for all x in that interval.

Manipulating Taylor series

You can obtain new Taylor series from standard Taylor series by:

- substituting for the variable
- adding, subtracting and multiplying Taylor series
- differentiating and integrating Taylor series term by term.

In each case, you can find an interval of validity for the new Taylor series from the known interval(s) of validity of the original Taylor series. In particular:

- A Taylor series about a point a for the sum, difference or product of two functions is valid for all values of x for which the Taylor series about a for both original functions are valid, and possibly for a larger interval of values.
- If an interval of validity for a Taylor series about a point a for a function f is an open interval, then it is also an interval of validity for the Taylor series about a for f' and for any antiderivative of f.

MST124 Unit 12 Complex numbers

Complex numbers and their arithmetic

The number i is defined to have the property $i^2 = -1$.

A **complex number** is a number of the form $a + bi$, where a and b are real numbers.

If $z = a + bi$, then

- the **real part** of z, denoted by $\operatorname{Re}(z)$, is a
- the **imaginary part** of z, denoted by $\operatorname{Im}(z)$, is b
- the **complex conjugate** of z, denoted by $\overline{z}$, is $a - bi$.

To manipulate complex numbers, use the usual rules of algebra and the fact that $i^2 = -1$.

To simplify a quotient of complex numbers, multiply the numerator and denominator by the complex conjugate of the denominator.

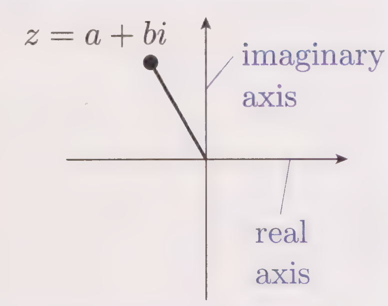
Square roots of a negative real number

If d is a positive real number, then the square roots of $-d$ are $\pm i\sqrt{d}$.

Properties of complex conjugates

$$\overline{z+w} = \overline{z} + \overline{w} \qquad \overline{z-w} = \overline{z} - \overline{w}$$
$$\overline{zw} = \overline{z}\,\overline{w} \qquad \overline{z/w} = \overline{z}/\overline{w}$$
$$z\,\overline{z} = a^2 + b^2, \quad \text{where } z = a + bi$$

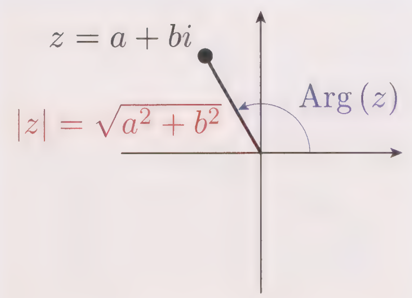
Modulus and argument



In the **complex plane** (the **Argand diagram**), the complex number $a + bi$ is represented by the point (a, b).

If $z = a + bi$, then:

- The **modulus** (**absolute value**) of z, denoted by $|z|$, is its distance from the origin, given by
$$|z| = \sqrt{a^2 + b^2}.$$
- An **argument** of z is an angle in radians measured anticlockwise from the positive real axis to the line between the origin and z.
- The **principal argument** of z, denoted by $\operatorname{Arg}(z)$, is the argument of z that lies in the interval $(-\pi, \pi]$.



The number 0 does not have an argument or a principal argument.

Properties involving modulus

$$z\overline{z} = |z|^2 \qquad |zw| = |z||w| \qquad \left|\frac{z}{w}\right| = \frac{|z|}{|w|} \qquad \frac{1}{z} = \frac{\overline{z}}{|z|^2}$$

Forms of complex numbers

Cartesian, polar and exponential form

A complex number z can be written in any of the following three forms.

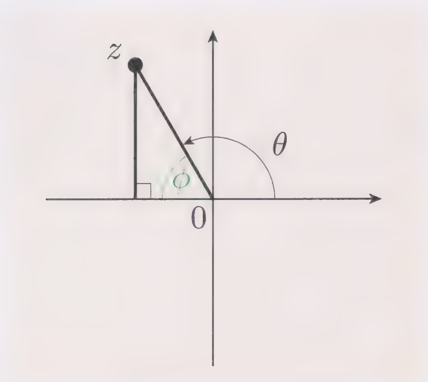
- **Cartesian form**: $z = a + bi$, where a and b are real numbers.
- **Polar form**: $z = r(\cos\theta + i\sin\theta)$, where r is the modulus of z and θ is an argument of z.
- **Exponential form**: $z = re^{i\theta}$, where r is the modulus of z and θ is an argument of z.

The expression $e^{i\theta}$ is defined by **Euler's formula**:

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

To convert from Cartesian form to polar or exponential form

Given a non-zero complex number $z = a + bi$, proceed as follows.



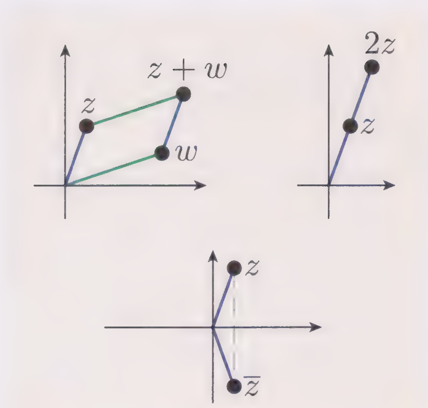
1. Find the modulus r, using $r = |z| = \sqrt{a^2 + b^2}$.
2. Find the principal argument θ:
   (a) Mark z in roughly the right position in the complex plane, draw the line from 0 to z, and mark and label the principal argument θ.
   (b) If z lies on one of the axes, then use your diagram to find the principal argument θ. Otherwise carry out steps (c) to (e).
   (c) Label the acute angle between the real axis and the line from 0 to z as ϕ.
   (d) Draw a line from z perpendicular to the real axis to form a right-angled triangle, and mark the lengths of the horizontal and vertical sides.
   (e) Use the triangle to work out the angle ϕ and hence the principal argument θ.
3. Write z in polar form $z = r(\cos\theta + i\sin\theta)$ or exponential form $z = re^{i\theta}$, as required.

To convert from polar or exponential form to Cartesian form

Write the complex number in polar form $z = r(\cos\theta + i\sin\theta)$, evaluate $\cos\theta$ and $\sin\theta$, and multiply out the brackets.

Euler's equation

$$e^{i\pi} + 1 = 0$$



Geometry of complex numbers

- Adding two complex numbers is similar to adding two vectors: a **parallelogram law** holds.
- Multiplying a complex number by a real number is similar to multiplying a vector by a scalar.
- Taking the complex conjugate of a complex number reflects it in the real axis.

Working with complex numbers in polar or exponential form

Complex conjugates

Polar form If $z = r(\cos\theta + i\sin\theta)$, then $\overline{z} = r(\cos(-\theta) + i\sin(-\theta))$.

Exponential form If $z = re^{i\theta}$, then $\overline{z} = re^{-i\theta}$.

Products and quotients

Polar form Let $z = r(\cos\theta + i\sin\theta)$ and $w = s(\cos\phi + i\sin\phi)$. Then

$$zw = rs\big(\cos(\theta+\phi) + i\sin(\theta+\phi)\big)$$
$$\frac{z}{w} = \frac{r}{s}\,(\cos(\theta-\phi) + i\sin(\theta-\phi)).$$

Exponential form Let $z = re^{i\theta}$ and $w = se^{i\phi}$. Then

$$zw = rs\,e^{i(\theta+\phi)}$$
$$\frac{z}{w} = \frac{r}{s}\,e^{i(\theta-\phi)}.$$

De Moivre's formula (formula for powers)

Polar form Let $z = r(\cos\theta + i\sin\theta)$. Then, for any integer n,

$$z^n = r^n(\cos n\theta + i\sin n\theta).$$

Exponential form Let $z = re^{i\theta}$. Then, for any integer n,

$$z^n = r^n e^{in\theta}.$$

Equality

If $r(\cos\theta + i\sin\theta) = s(\cos\phi + i\sin\phi)$ or, equivalently, $re^{i\theta} = se^{i\phi}$, then

$$r = s \quad \text{and} \quad \theta = \phi + 2m\pi \text{ for some integer } m.$$

To work out a power z^n of a complex number z in Cartesian form

1. Write z in polar form, $z = r(\cos\theta + i\sin\theta)$.
2. Apply de Moivre's formula.
3. Convert the result back to Cartesian form.

(For small positive values of n this strategy is unnecessary – use the usual arithmetic of complex numbers.)

Some formulas useful for finding trigonometric identities

Special case of de Moivre's formula

$$\cos n\theta + i\sin n\theta = (\cos\theta + i\sin\theta)^n$$

Formulas deduced from Euler's formula

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

Polynomial equations

A **polynomial equation** is an equation of the form 'polynomial expression = 0', where the polynomial expression has degree at least 1.

A **complex solution** of a polynomial equation is a solution of the equation that is a complex number. (It may be a real number, since every real number is also a complex number.)

To find the complex solutions of a quadratic equation

Simplify it, then use one of the following methods.

Completing the square

As on page 26, but if necessary use the fact that if d is a positive real number, then the two square roots of $-d$ are $\pm i\sqrt{d}$.

The quadratic formula

As on page 26, the solutions of the quadratic equation $az^2 + bz + c = 0$, where a, b and c are real numbers, are given by

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

If $b^2 - 4ac$ is negative, say $b^2 - 4ac = -d$ where d is positive, then $\pm\sqrt{b^2 - 4ac}$ means $\pm i\sqrt{d}$ (the two square roots of the negative number $b^2 - 4ac$).

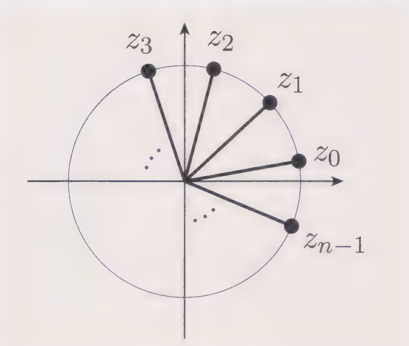
Roots of complex numbers

An ***n*th root** of a complex number a is a solution z of the equation $z^n = a$, where a is a complex number and n is a positive integer.

An ***n*th root of unity** is a solution z of the equation $z^n = 1$, where n is a positive integer.

To find the complex solutions of the equation $z^n = a$, where $a \neq 0$

(There are n solutions, equally spaced on a circle centred at 0.)



1. Write the unknown z in polar form, in terms of an unknown modulus r and an unknown argument θ, and write the number a in polar form.
2. Substitute the polar forms of z and a into the equation, and apply de Moivre's formula to find the polar form of the left-hand side.
3. Compare moduli to find the value of r.
4. Compare arguments to find n successive possible values of θ.
5. Hence write down the n possible values of z.

It is usually convenient to use arguments in the interval $[0, 2\pi)$.

The fundamental theorem of algebra

Every polynomial $a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ of degree $n \geq 1$ has a factorisation $a_n(z - z_1)(z - z_2)\cdots(z - z_n)$, where $z_1, z_2, \ldots, z_n$ are complex numbers, some of which may be equal to others.

MST125

MST125 Unit 3 Number theory

Factors and multiples

An integer a is **divisible** by a non-zero integer n if there is a third integer k such that $a = kn$.

If a, n and k are integers such that $a = kn$, then a is a **multiple** of n, and n is a **factor** or **divisor** of a.

The **highest common factor** (HCF), or **greatest common divisor**, of two integers a and b, not both zero, is the greatest positive integer n that is a factor of both a and b.

The division theorem

Suppose that a is an integer and n is a positive integer. Then there are unique integers q and r such that

$$a = qn + r \quad \text{and} \quad 0 \le r < n.$$

The **quotient** of the division of a by n is the integer q.

The **remainder** of the division is the integer r.

Euclid's algorithm

To find the highest common factor of positive integers a and b, where $a > b$

1. By applying the division theorem repeatedly, form a list of equations:

$$\begin{aligned} a &= q_1 b + r_1 \\ b &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ r_2 &= q_4 r_3 + r_4 \\ r_3 &= q_5 r_4 + r_5 \\ &\;\;\vdots \qquad\;\; \vdots \end{aligned}$$

2. Stop when you obtain an equation in which the remainder is 0.
3. The highest common factor of a and b is the remainder in the second-to-last equation.

Bézout's identity and backwards substitution

Bézout's identity

Suppose that a and b are integers, not both zero, and let d be their highest common factor. Then there are integers v and w such that

$$av + bw = d.$$

To find v and w in Bézout's identity for given positive integers a and b

First apply Euclid's algorithm to find the highest common factor d of a and b.

Then proceed as follows, applying **backwards substitution**.

1. Cross out or ignore the last equation from Euclid's algorithm (the one with remainder zero).
2. In each remaining equation, circle each number except the quotient. For the rest of this strategy, treat the circled numbers like variables: do not combine them with other numbers.
3. Rearrange each equation to make the remainder the subject. In particular, the bottom equation then gives an expression for d.
4. Substitute one of the circled numbers in the expression for d using the second-bottom equation. Simplify to obtain a new expression for d.
5. Substitute one of the circled numbers in the new expression for d using the third-bottom equation. Simplify to obtain another new expression for d.
6. Continue in this way, working upwards through all the equations.
7. The final expression for d gives the required values of v and w.

To find v and w in Bézout's identity for given integers a and b, where at least one of a and b is negative

Apply the procedure above to $|a|$ and $|b|$, which have the same highest common factor as a and b. Adjust the signs in the final expression for d to obtain an expression that gives the required values of v and w.

Congruences

Two integers a and b are **congruent modulo** n if they each have the same remainder on division by n, where n is a positive integer. We write this as $a \equiv b \pmod{n}$ and call a statement of this form a **congruence**.

For any integers a and b the following statements are equivalent:

- a and b are congruent modulo n
- $a - b$ is divisible by n
- there is an integer k such that $a = b + nk$.

Modular arithmetic is arithmetic carried out with congruences.

Three properties of congruences

Reflexivity $a \equiv a \pmod{n}$, for any a.

Symmetry If $a \equiv b \pmod{n}$ then $b \equiv a \pmod{n}$.

Transitivity If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$.

Residues

The **residue class**, or **congruence class**, of a **modulo** n is the set of all integers congruent to a modulo n.

The **least residue** of a **modulo** n is the smallest number greater than or equal to zero in the residue class of a modulo n. It is the remainder obtained when a is divided by n.

Rules for combining congruences

If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, and m is a positive integer, then

$$a + c \equiv b + d \pmod{n} \qquad ac \equiv bd \pmod{n}$$

$$a - c \equiv b - d \pmod{n} \qquad a^m \equiv b^m \pmod{n}.$$

Fermat's little theorem

Let p be a prime number, and let a be an integer that is not a multiple of p. Then

$$a^{p-1} \equiv 1 \pmod{p}.$$

Divisibility tests

The **digit sum** of an integer is the sum of its digits.

If an integer is divisible by 3, then its digit sum is divisible by 3, and vice versa.

If an integer is divisible by 9, then its digit sum is divisible by 9, and vice versa.

Multiplicative inverses

A **multiplicative inverse** of a modulo n is an integer v such that

$$av \equiv 1 \pmod{n}.$$

Two integers are **coprime** if their highest common factor is 1.

An integer a has a multiplicative inverse modulo n if a and n are coprime. Otherwise a has no multiplicative inverse modulo n.

- If a has a multiplicative inverse modulo n, then every number congruent to the multiplicative inverse modulo n is also a multiplicative inverse of a modulo n.
- If $a \equiv b \pmod{n}$, then every multiplicative inverse of a modulo n is also a multiplicative inverse of b modulo n.

To check whether a has a multiplicative inverse modulo n (where $0 < a < n$) and find one if it does

1. Check whether a and n are coprime. You might do this by inspection, by using Euclid's algorithm, or by another means.
2. If a and n are not coprime, then there is no multiplicative inverse.
3. If a and n are coprime, then proceed as follows.
   - If n is small (say $n \leq 13$), then try the values 1, 2, 3, …, $n-1$ in turn until you find a multiplicative inverse.
   - If n is large (say $n \geq 14$), then use Euclid's algorithm (which you may already have applied in step 1) and backwards substitution to find an equation of the form $av + nw = 1$; that is, $av = 1 - nw$. Then v is a multiplicative inverse of a modulo n.

Linear congruences

A **linear congruence** is a congruence of the form $ax \equiv b \pmod{n}$, where a, b and n are known, and x is unknown.

A **solution** of a linear congruence $ax \equiv b \pmod{n}$ is a value of x that satisfies the congruence. If c is a solution of a linear congruence, then every number congruent to c modulo n is also a solution. So the solutions are given by linear congruences of the form $x \equiv c \pmod{n}$.

Solving a linear congruence means finding its solutions.

Solutions of the linear congruence $ax \equiv b \pmod{n}$

Let d be the highest common factor of a and n.

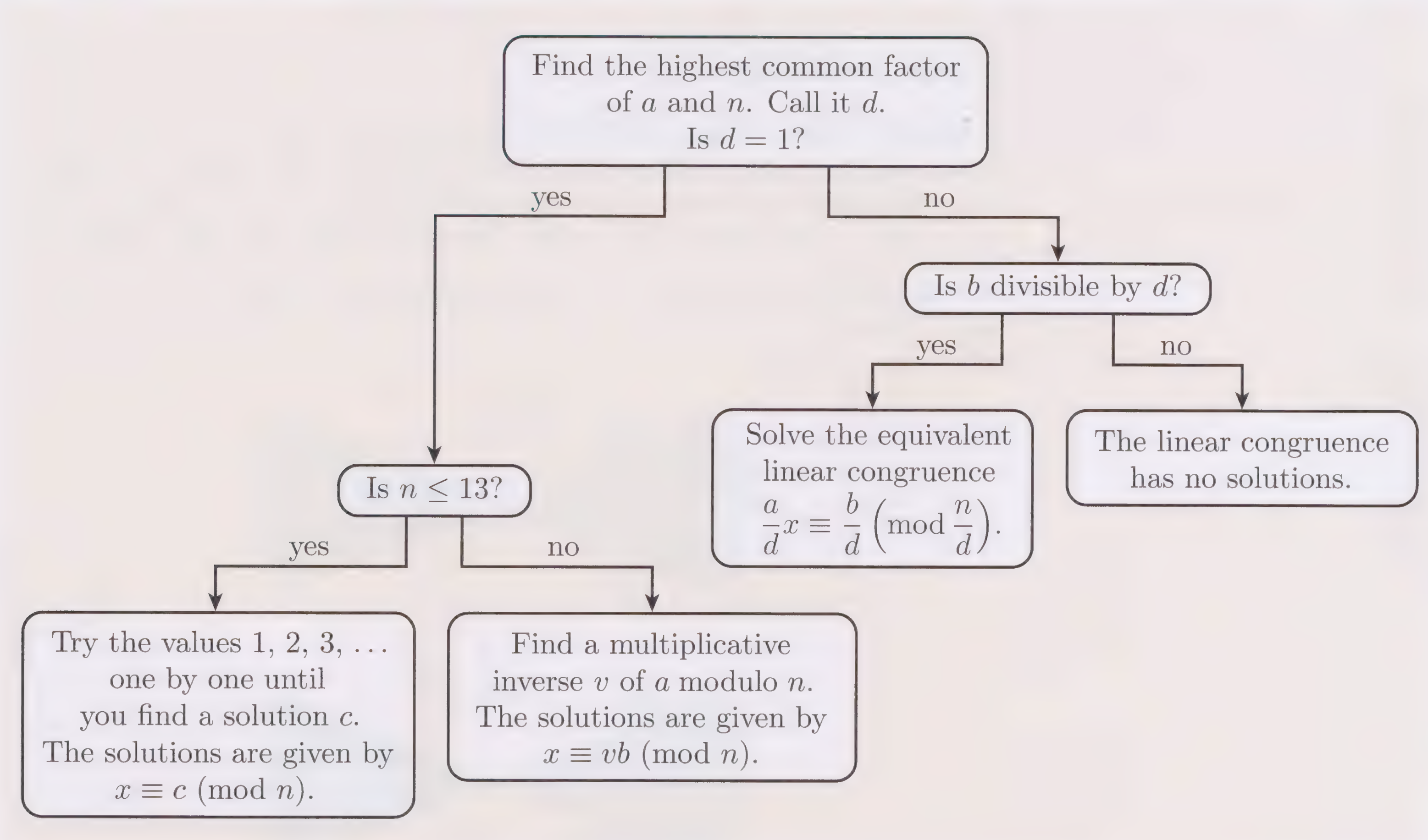
- If $d = 1$, then the linear congruence has solutions, which are given by

  $$x \equiv vb \pmod{n},$$

  where v is any multiplicative inverse of a modulo n.
- If b is not divisible by d, then the linear congruence has no solutions.
- If b is divisible by d, then the linear congruence has solutions, which are the same as the solutions of the linear congruence

  $$\frac{a}{d}x \equiv \frac{b}{d} \left(\bmod \frac{n}{d}\right).$$

Decision tree for solving the linear congruence $ax \equiv b \pmod{n}$



Ciphers

A **cipher** is an algorithm used to transform a message to disguise its content.

Enciphering is the process of applying a cipher to a message. **Deciphering** is the process of recovering the original message from an enciphered message.

An **affine cipher** is a rule E for enciphering the integers $0, 1, 2, \ldots, 25$ given by

$$E(x) \equiv ax + b \pmod{26},$$

where a and b are integers, and a and 26 are coprime.

(More generally, the number 26 here can be replaced by any integer $n \geq 2$ to give a rule for enciphering the integers $0, 1, 2, \ldots, n-1$.)

Deciphering rule for an affine cipher

The affine cipher E with rule

$$E(x) \equiv ax + b \pmod{26}$$

has deciphering rule

$$D(y) \equiv v(y - b) \pmod{26},$$

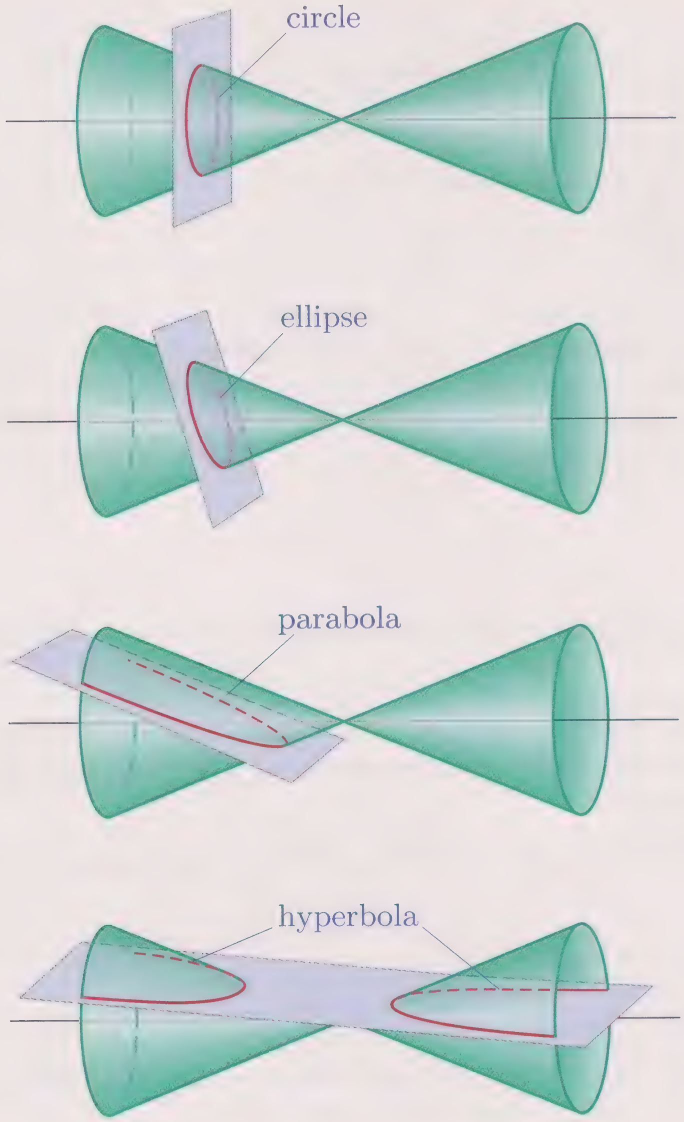
where v is any multiplicative inverse of a modulo 26.

MST125 Unit 4 Conics

Conic sections

A **double cone** consists of two identical, infinitely long, hollow circular cones joined at their vertices and aligned symmetrically along a straight line. The **apex** or **vertex** of the double cone is the point at which the two cones meet. The **axis** is the straight line that passes centrally through it.

A **conic section**, or **conic**, is a curve obtained by slicing a double cone with a plane.



Degenerate conics

A **degenerate conic section**, or **degenerate conic**, is obtained when the slicing plane passes through the apex of the double cone. It may be a single point, a straight line or two intersecting straight lines.

A **non-degenerate conic section**, or **non-degenerate conic**, is a conic section that is not degenerate.

Focus–directrix definitions of conics

If F is a point, d is a line that does not pass through F, and e is a positive number, then the curve formed by the points P satisfying the equation

$$PF = e\,Pd$$

is

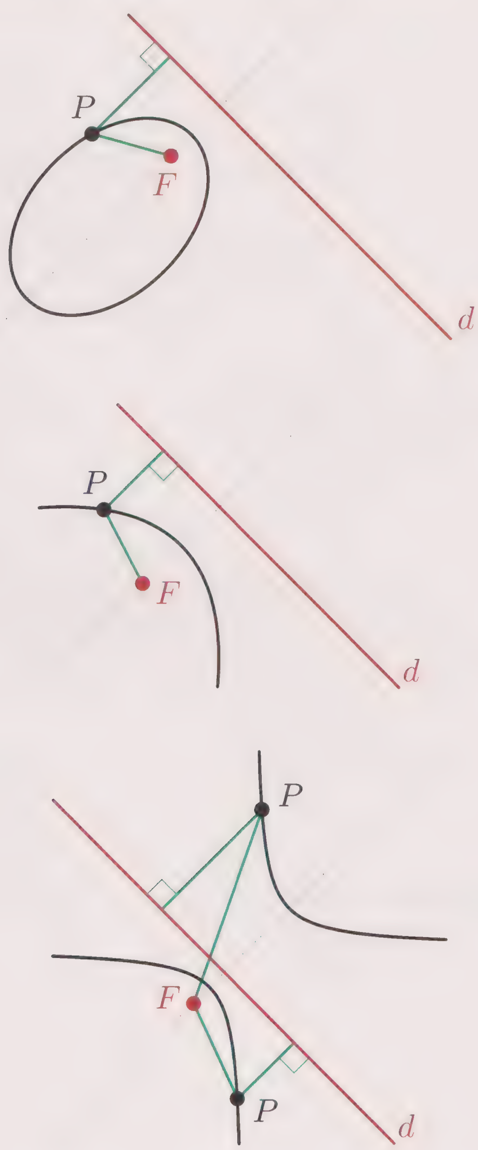
- an **ellipse** if $0 < e < 1$
- a **parabola** if $e = 1$
- a **hyperbola** if $e > 1$.

Here PF is the distance between P and F, and Pd is the perpendicular distance between P and d.

The point F is called the **focus**, the line d is called the **directrix**, and the number e is called the **eccentricity**.

A circle does not have a focus–directrix property.

The plural of 'focus' is 'foci' and the plural of 'directrix' is 'directrices'.



ellipse ($0 < e < 1$)

parabola ($e = 1$)

hyperbola ($e > 1$)

Terminology for ellipses and hyperbolas

The **centre** of an ellipse or hyperbola is the point where its axes of symmetry cross. The asymptotes of a hyperbola cross at its centre.

The **vertices** of an ellipse or hyperbola are the points where it crosses its axes of symmetry.

An **axis** of an ellipse is a line segment that joins two vertices on the same axis of symmetry of the ellipse. The **major axis** is the longer axis and the **minor axis** is the shorter axis.

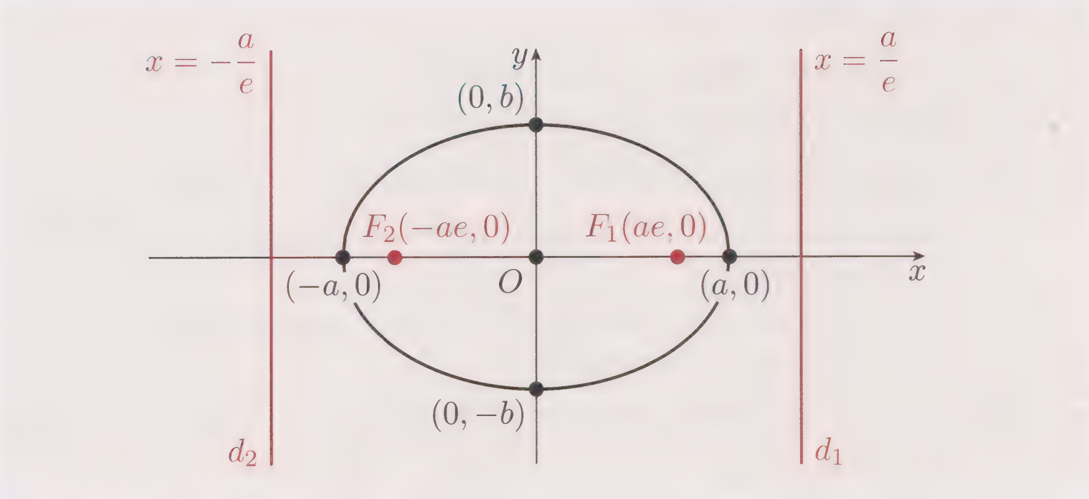
The **branches** of a hyperbola are its two continuous parts.

A **rectangular hyperbola** is a hyperbola whose asymptotes are perpendicular.

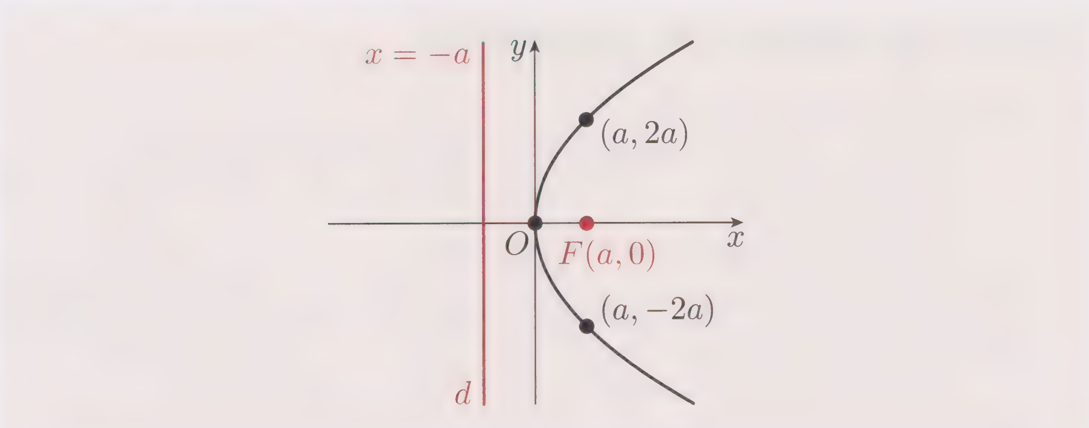
Properties of ellipses, parabolas and hyperbolas in standard position

	Ellipse	Parabola	Hyperbola
Equation	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$	$y^2 = 4ax$	$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$
	(where $a > b > 0$)	(where $a > 0$)	(where $a, b > 0$)
Vertices	$(\pm a, 0)$, $(0, \pm b)$	$(0, 0)$	$(\pm a, 0)$
Asymptotes			$y = \pm \dfrac{b}{a} x$
Eccentricity	$0 < e < 1$	$e = 1$	$e > 1$
	$e = \sqrt{1 - \dfrac{b^2}{a^2}}$		$e = \sqrt{1 + \dfrac{b^2}{a^2}}$
Foci	$(\pm ae, 0)$	$(a, 0)$	$(\pm ae, 0)$
Directrices	$x = \pm \dfrac{a}{e}$	$x = -a$	$x = \pm \dfrac{a}{e}$
Axes of symmetry	x-axis, y-axis	x-axis	x-axis, y-axis
Standard parametrisation	$x = a\cos t$, $y = b\sin t$ $(0 \leq t \leq 2\pi)$	$x = at^2$, $y = 2at$	$x = a\sec t$, $y = b\tan t$ $\left(-\dfrac{\pi}{2} < t < \dfrac{\pi}{2}, \dfrac{\pi}{2} < t < \dfrac{3\pi}{2}\right)$

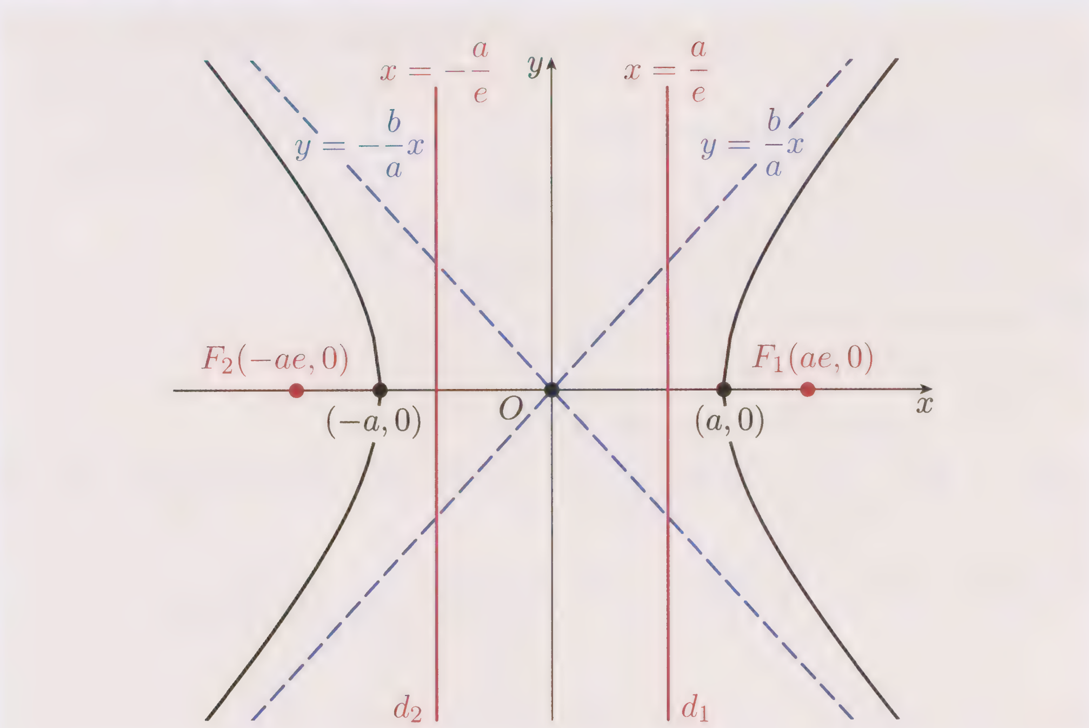
Ellipse in standard position



Parabola in standard position



Hyperbola in standard position



Further properties of ellipses

- When the eccentricity e is
  - close to 0, the ellipse is nearly circular
  - close to 1, the ellipse is very squashed.
- The ellipse with equation $x^2/a^2 + y^2/b^2 = 1$ can be obtained
  - from the circle with equation $x^2 + y^2 = a^2$ by scaling it vertically by the factor b/a
  - from the circle with equation $x^2 + y^2 = b^2$ by scaling it horizontally by the factor a/b.
- If the point P lies on the ellipse with equation $x^2/a^2 + y^2/b^2 = 1$, then

  $$PF_1 + PF_2 = 2a,$$

  where F_1 and F_2 are the foci of the ellipse.

Further properties of hyperbolas

- When the eccentricity e is
  - close to 1, the hyperbola is very squashed (each branch is closely 'folded')
  - large, the hyperbola is stretched out.
- A hyperbola with an equation of the form $x^2 - y^2 = a^2$ is rectangular, with asymptotes $y = \pm x$.

Conics not in standard position

General equation for a conic

Every conic, no matter what its position in the plane, has an equation of the form

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0,$$

where A, B, C, D, E and F are constants.

(Not every equation of this form represents a conic.)

To determine the type of a conic

Suppose that the equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0,$$

where A, B, C, D, E and F are constants, represents a non-degenerate conic.

- If $B^2 - 4AC < 0$, then the conic is an ellipse or a circle.
  (It is a circle if $A = C$ and $B = 0$, and an ellipse otherwise.)
- If $B^2 - 4AC = 0$, then the conic is a parabola.
- If $B^2 - 4AC > 0$, then the conic is a hyperbola.

Parametric equations

A pair of **parametric equations** for a line or curve is a pair of equations that express the x- and y-coordinates of any point on the line or curve in terms of a third variable, the **parameter**, usually denoted by t.

A **parametrisation** of a line or curve is any way of describing the line or curve in terms of a parameter. For example, a parametrisation may be given by a pair of parametric equations, or by an expression for the coordinates of each point on the line or curve in terms of a parameter.

A **parametrisation** of part of a line or curve is any way of describing the part of the line or curve in terms of a parameter. It is usually necessary to include a restricted range of values for the parameter.

Parametric equations for a straight line

Every pair of parametric equations of the form

$$x = at + b, \quad y = ct + d,$$

where a, b, c and d are constants with a and c not both zero, describes a straight line.

Parametric equations for a straight line, arising from a point on it and its gradient

The line that passes through the point (x_0, y_0) and has gradient m has parametric equations

$$x = x_0 + t, \quad y = y_0 + mt.$$

With this parametrisation, $t = 0$ corresponds to the point (x_0, y_0).

Parametric equations for a straight line, arising from two points on it

The line that passes through the points (x_0, y_0) and (x_1, y_1) has parametric equations

$$x = x_0 + t(x_1 - x_0), \quad y = y_0 + t(y_1 - y_0).$$

With this parametrisation, $t = 0$ corresponds to the point (x_0, y_0) and $t = 1$ corresponds to the point (x_1, y_1).

Parametric equations for a circle

The circle with centre (p, q) and radius r has parametric equations

$$x = p + r\cos t, \quad y = q + r\sin t.$$

Parametric equations for a translated curve

The curve obtained by translating the curve with parametric equations

$$x = f(t), \quad y = g(t),$$

by p units right and q units up (where p and q can be positive, negative or zero) has parametric equations

$$x = p + f(t), \quad y = q + g(t).$$

MST125 Unit 5 Statics

Mass and acceleration

The **mass** of an object is a measure of the amount of matter that constitutes it. The SI unit for mass is the **kilogram** (kg).

The **acceleration** of an object is its rate of change of velocity. It is a vector quantity. The SI units for acceleration are **metres per second per second**, also called **metres per second squared** ($\mathrm{m\,s^{-2}}$).

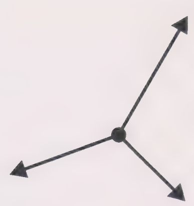
The **acceleration due to gravity** is the acceleration with which objects fall when acted on by gravity alone. Its magnitude (on Earth) is $g = 9.8\,\mathrm{m\,s^{-2}}$ (to 2 s.f.) and its direction is vertically downwards.

Force

A **force** is an influence that can cause an object to accelerate. It is a vector quantity.

One **newton** (N) is the magnitude of force that, when acting on an object of mass 1 kg, causes the object to accelerate at 1 metre per second per second ($1\,\mathrm{m\,s^{-2}}$) in the direction of the force.

The **resultant force**, or **net force**, on an object is the vector sum of the forces that act on it.



A **force diagram** represents the forces that act on an object. The object is represented by a dot. The forces are represented by arrows whose lengths and directions roughly represent the magnitudes and directions of the forces.

Newton's first and third laws of motion

I An object remains at rest or continues to move with a constant velocity, unless acted on by a resultant force.

III For each force exerted by one object on a second object, there is a force of equal magnitude in the opposite direction, exerted by the second object on the first.

An **action** and an **equal and opposite reaction** are two forces as described in Newton's third law.

Equilibrium

An object is in **equilibrium** if the sum of the forces that act on it is $\mathbf{0}$.

It is **at rest** if its position does not change.

It is in **static equilibrium** if it is in equilibrium and remains at rest.

Equilibrium condition for a particle

If an object modelled by a particle is acted on by the n forces $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$, and is in equilibrium, then

$$\mathbf{F}_1 + \mathbf{F}_2 + \cdots + \mathbf{F}_n = \mathbf{0}.$$

Modelling assumptions

A **modelling assumption** is a simplification made to a real-world situation to obtain a model that can be dealt with mathematically.

A **particle** is an object that has mass but no size, and so occupies a single point in space.

A **model string** is a string (or rope, wire, chain, and so on) that is **inextensible** (does not stretch) and **light** (has no mass).

A **smooth** surface imposes no resistance to motion parallel to itself; that is, it is frictionless. A **rough** surface is one that is not smooth.

A **model pulley** is an object with no mass or size, over which a model string may pass without any resistance to motion. The magnitude of the tension in a model string passing over a model pulley is the same on each side of the model pulley.

Weight, tension and normal reaction

The **weight** of an object is a force that acts on the object and is directed vertically downwards. If the object has mass m (in kg), then its weight has magnitude mg (in N), where $g = 9.8\,\mathrm{m\,s^{-2}}$.

A **tension** in a taut string (or rope, wire, chain, and so on) is a force that pulls at a point of attachment of the string, along the line of the string. Its magnitude adjusts to balance other forces, up to some limit.

The **normal reaction** of a surface on an object resting on the surface is a force that pushes on the object, at right angles to the surface. Its magnitude adjusts to balance other forces, up to some limit.

Static friction

A **static friction force** from a flat surface on an object at rest on the surface is a force that acts on the object, in the direction parallel to the surface and opposite to any possible motion of the object across the surface. Its magnitude adjusts to balance other forces, up to its maximum magnitude, which is

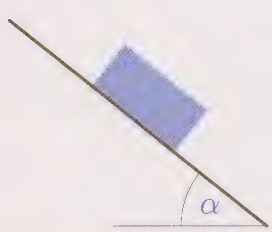
$$\mu|\mathbf{N}|,$$

where μ is the **coefficient of static friction** (a constant that depends on the two materials in contact), and $\mathbf{N}$ is the normal reaction of the surface on the object.

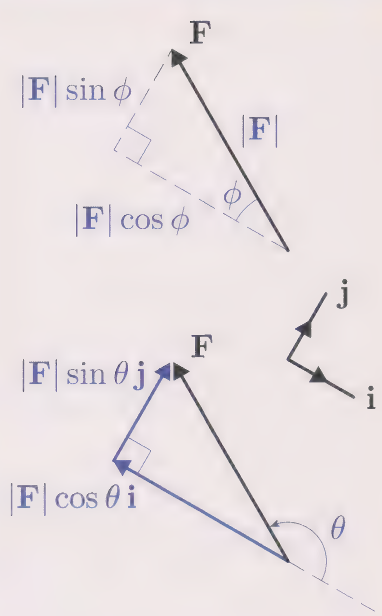
The object is **on the point of slipping**, or **held by limiting friction**, if the magnitude of the friction force is exactly at its maximum value $\mu|\mathbf{N}|$.

Coefficient of static friction in terms of angle of inclination

If an object rests on a plane inclined at the angle α, and is on the point of slipping, then the coefficient μ of static friction between the object and the plane is given by $\mu = \tan\alpha$.



Resolving forces



Resolving a force in the **i**- and **j**-directions means finding its components in the **i**- and **j**-directions, where **i** and **j** are the Cartesian unit vectors.

To resolve a force F in the i- and j-directions

Sketch **F** as the hypotenuse of a right-angled triangle whose other two sides are parallel to **i** and **j**. Find an acute angle in this triangle, and hence find the magnitudes of the components. Use the directions of **i** and **j** to find the signs of the components.

Alternatively, use the following formula. If **F** makes the angle θ with the **i**-direction (where θ is measured anticlockwise from the **i**-direction), then

$$\mathbf{F} = |\mathbf{F}| \cos\theta\, \mathbf{i} + |\mathbf{F}| \sin\theta\, \mathbf{j}.$$

To find the magnitude and direction of a force from its components

The force $\mathbf{F} = a\mathbf{i} + b\mathbf{j}$ has magnitude $|\mathbf{F}| = \sqrt{a^2 + b^2}$.

To find the direction of a force $\mathbf{F} = a\mathbf{i} + b\mathbf{j}$, sketch **F** as the hypotenuse of a right-angled triangle whose other two sides are parallel to **i** and **j**. Find an acute angle in this triangle and hence find the direction of **F**. Describe the direction relative to the horizontal, the vertical or some direction intrinsic to the situation.

Resolving a vector equation

Resolving a vector equation into components means considering the components separately, to obtain simultaneous scalar equations.

To solve a statics problem

1. Make any necessary modelling assumptions.
2. Draw a diagram of the physical situation, annotating it with any relevant information.
3. Identify all the forces acting and draw a separate force diagram, stating any known magnitudes of forces.
4. Choose directions for **i** and **j** and draw them on the force diagram.
5. Express the forces in component form, in terms of unknown quantities where necessary.
6. Use the equilibrium condition and any other appropriate laws to obtain one or more equations.
7. Solve the equations.
8. State a conclusion.

Choosing the directions of i and j strategically

- If a problem involves two perpendicular forces, consider choosing **i** and **j** to be parallel to these forces.
- If a problem involves finding the magnitude of a force whose direction you know, consider choosing **i** or **j** to be parallel to this force.

MST125 Unit 6 Geometric transformations

Transformations of the plane

The **plane** is denoted by $\mathbb{R}^2$. Its elements are points, denoted by coordinates of the form (x, y), or by column vectors of the form $\begin{pmatrix} x \\ y \end{pmatrix}$.

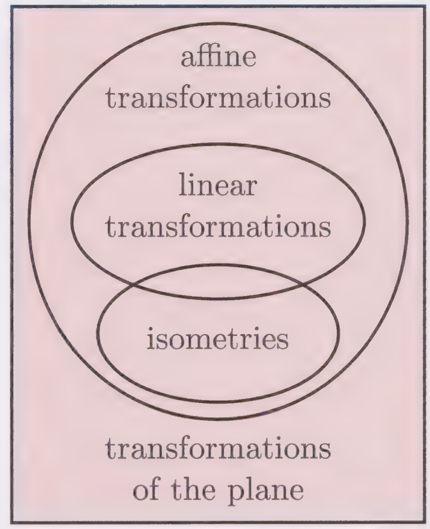
A **transformation of the plane** is a function whose domain and codomain are subsets of $\mathbb{R}^2$.

In MST125 the word 'transformation' is assumed to mean 'transformation of the plane', and it is assumed that the codomain of every transformation is $\mathbb{R}^2$. The domain convention (see page 29) applies.

The **identity transformation** leaves all points in the plane fixed.

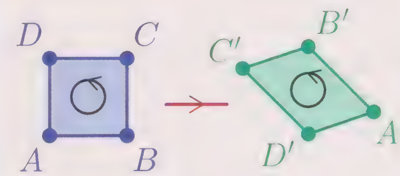
The **zero transformation** maps all points in the plane to the origin.

Types of transformation include *isometries*, *linear transformations* and *affine transformations*.

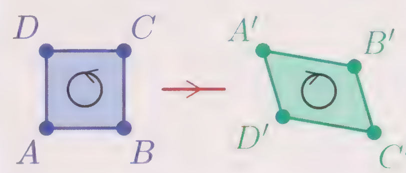


A transformation **preserves orientation** if it leaves the cyclical order of the vertices of each figure unchanged.

A transformation **reverses orientation** if it reverses the cyclical order of the vertices of each figure.



Preserving orientation



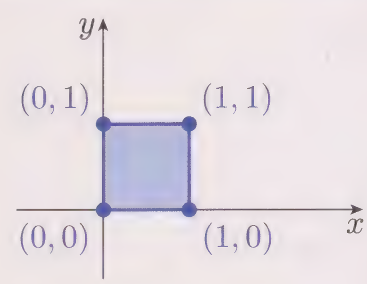
Reversing orientation

Plane figures

A **plane figure**, or just **figure**, is a subset of $\mathbb{R}^2$.

A **polygon** is a plane figure whose boundary consists of line segments.

The **unit square** is the square with vertices $(0, 0)$, $(1, 0)$, $(1, 1)$ and $(0, 1)$.



The unit square

The **image** $f(A)$ of a figure A under a transformation f is the figure whose points are the images under f of the points in A. We say that f **maps** A to $f(A)$.

Composite and inverse transformations

- The **composite transformation** $g \circ f$ formed from the transformation f followed by the transformation g is the transformation whose rule is

  $$(g \circ f)(x, y) = g(f(x, y)),$$

  and whose domain consists of all the points (x, y) in the domain of f such that $f(x, y)$ is in the domain of g.

  Composition of transformations is not commutative; that is, the transformation $g \circ f$ need not be the same as the transformation $f \circ g$.

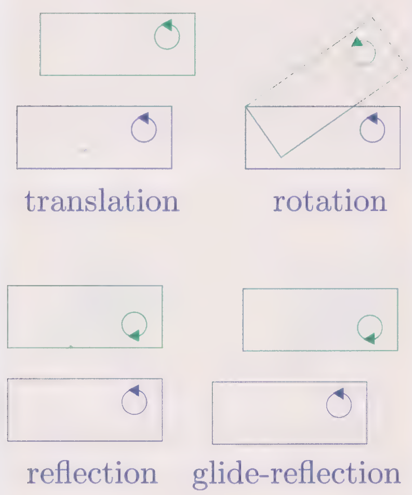
- If a transformation f is one-to-one (that is, invertible), then its **inverse transformation** f^{-1} is the transformation such that if f maps the point (x, y) to (x', y'), then f^{-1} maps the point (x', y') back to (x, y). The domain of f^{-1} is the image set of f, and the image set of f^{-1} is the domain of f.

Isometries

An **isometry** is a transformation f such that the distance between any two points P and Q is equal to the distance between their images $f(P)$ and $f(Q)$.

Types of isometry

Every isometry is one of the following four types.



- A **translation** p units horizontally and q units vertically.
  This maps each point (x, y) to the point $(x + p, y + q)$. The numbers p and q can be positive, negative or zero.

  The **associated vector** of such a translation is the displacement vector $p\mathbf{i} + q\mathbf{j}$, where as usual $\mathbf{i}$ and $\mathbf{j}$ are the Cartesian unit vectors in the positive directions of the x- and y-axes, respectively.
- A **rotation** through an angle θ about a point C.
  This rotates each point through the angle θ about C. (As usual, the rotation is anticlockwise if θ is positive, and clockwise if θ is negative.) The point C is the **centre of rotation**.
- A **reflection** in a line ℓ.
  This maps each point P to the point P' on the other side of ℓ in such a way that ℓ is the perpendicular bisector of the line segment PP'. Points on ℓ remain fixed. The line ℓ is the **line of reflection**.
- A **glide-reflection** in a line ℓ.
  This is a reflection in ℓ followed by a translation parallel to ℓ. (Or vice versa; the order of composition does not matter here.)

Properties of isometries

- An isometry maps any polygon to a polygon of the same size and shape. It maps the vertices of the polygon to the vertices of the image polygon.
- Translations and rotations preserve orientation.
  Reflections and glide-reflections reverse orientation.
- A composite of isometries is an isometry.
- Every isometry is invertible and its inverse is an isometry.
- A **fixed point** of an isometry is a point that the isometry maps to itself. A translation and a glide-reflection have no fixed points, a rotation has one fixed point (the centre of rotation) and a reflection has a line of fixed points (the line of reflection).

Symmetries

A **symmetry** of a pattern or figure is an isometry that leaves the appearance of the pattern or figure unchanged.

A composite of symmetries of a pattern or figure is a symmetry of the pattern or figure.

The inverse of a symmetry of a pattern or figure is a symmetry of the pattern or figure.

Composing isometries

Two translations The composite (in either order) of the two translations $f(x,y) = (x+a, y+b)$ and $g(x,y) = (x+c, y+d)$ is the translation $(f \circ g)(x,y) = (g \circ f)(x,y) = (x+a+c, y+b+d)$. That is, its associated vector is the sum of the associated vectors of f and g.

Two rotations about the same point The composite (in either order) of the rotation through the angle θ about the point P and the rotation through the angle ϕ about the same point P is the rotation through the angle $\theta + \phi$ about P.

Two reflections in lines that meet The composite transformation formed from the reflection in the line ℓ followed by the reflection in the intersecting line m is the rotation about the point of intersection of ℓ and m through twice the angle from ℓ to m.

Two reflections in parallel lines The composite transformation formed from the reflection in the line ℓ followed by the reflection in the parallel line m is the translation through twice the distance from ℓ to m, in the direction from ℓ to m perpendicular to ℓ and m.

A reflection and a rotation about a point on the line of reflection The composite transformation formed from the reflection in the line ℓ followed by the rotation through the angle ϕ about the point P on ℓ is the reflection in the line obtained by rotating ℓ through $\phi/2$ about P.

The composite transformation formed from these transformations in the opposite order is the reflection in the line obtained by rotating ℓ through $-\phi/2$ about P.

A reflection and a rotation A composite of a reflection and a rotation is usually a glide-reflection, but sometimes a reflection.

A reflection and a glide-reflection A composite of a reflection and a glide-reflection is a translation or a rotation.

Inverting isometries

Translations The inverse of the translation $f(x,y) = (x+a, y+b)$ is the translation $f^{-1}(x,y) = (x-a, y-b)$. That is, its associated vector is the negative of the associated vector of f.

Rotations The inverse of the rotation through the angle θ about a point is the rotation through the angle $-\theta$ about the same point.

Reflections Every reflection is **self-inverse**; that is, it is its own inverse.

Conjugation using an isometry

Conjugation involves composing a transformation h with an isometry g and its inverse g^{-1}, to give the transformation $g^{-1} \circ h \circ g$. This transformation performs the same action as the transformation h but at a different location of the plane, as determined by g.

Linear transformations

A **linear transformation** is a transformation of the form

$$f(\mathbf{x}) = \mathbf{A}\mathbf{x},$$

where $\mathbf{A}$ is a 2×2 matrix.

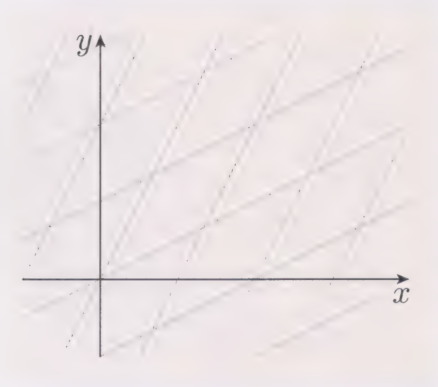
We say that f is **represented** by $\mathbf{A}$, and that $\mathbf{A}$ is the **matrix of** f.

To find the matrix of a linear transformation

If f is a linear transformation, and $f(1,0) = (a,c)$ and $f(0,1) = (b,d)$, then the matrix of f is $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Properties of linear transformations

- A linear transformation leaves the origin fixed.
- Every isometry that fixes the origin is a linear transformation.



Grids

A **grid** consists of two families of parallel lines that partition the plane into congruent parallelograms (which may be rectangles or squares) called **grid cells**. A grid is **rectangular** if its two families of lines are perpendicular, and **skewed** otherwise.

In the **unit grid**, the families of lines consist of the x- and y-axes and the lines parallel to them, one unit apart. Its grid cells are congruent to the unit square.

Properties of one-to-one linear transformations

A one-to-one linear transformation $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$

- **preserves linearity** (maps lines to lines)
- **preserves parallelism** (maps parallel lines to parallel lines)
- maps the unit grid to a grid whose grid cells include and are congruent to the image of the unit square under f
- preserves orientation if $\det \mathbf{A}$ is positive;
  reverses orientation if $\det \mathbf{A}$ is negative
- scales areas by the factor $|\det \mathbf{A}|$.

Linear transformations that are not one-to-one

A **flattening** is a linear (or affine) transformation whose image set is a line or a single point. The following are equivalent criteria for a linear transformation $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ to be a flattening:

- one column of $\mathbf{A}$ is a scalar multiple (possibly the zero multiple) of the other
- $\det \mathbf{A} = 0$
- f is not one-to-one.

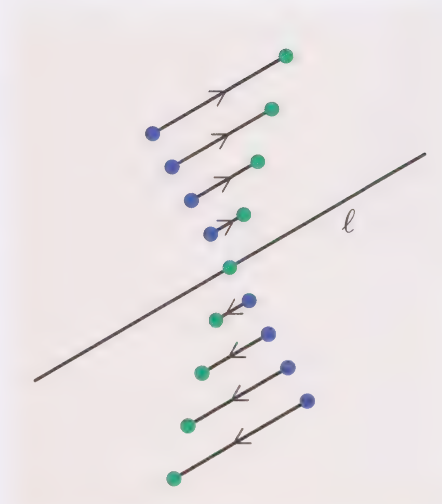
Scalings

The (k,l)-**scaling**, where $k \neq 0$ and $l \neq 0$, is the transformation that multiplies the horizontal displacement of each point from the y-axis by k, and multiplies the vertical displacement of each point from the x-axis by l. The numbers k and l are the **horizontal scale factor** and the **vertical scale factor**, respectively.

The k-**dilation** is the (k,k)-scaling, where $k \neq 0$. The number k is the **scale factor**.

Shears

The **shear** with **shear factor** k about the line ℓ is the transformation that displaces each point P in a direction parallel to ℓ, 'clockwise' if the shear factor is positive and 'anticlockwise' if it is negative, by the distance kd where d is the perpendicular distance of P from ℓ.



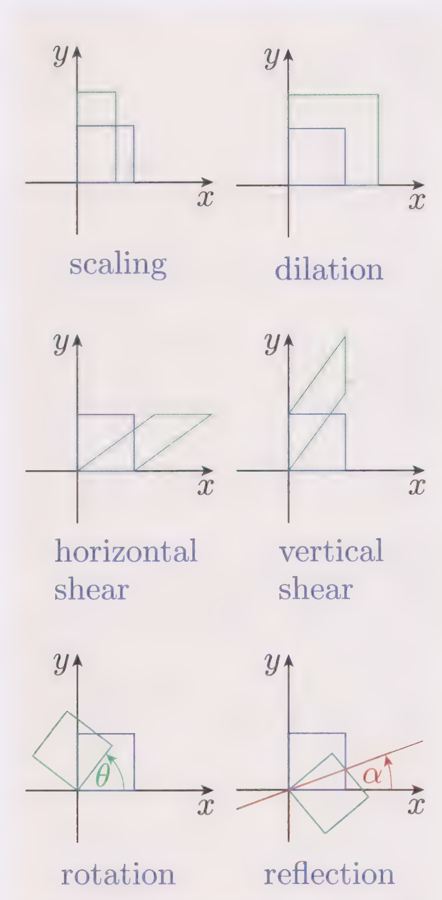
A **horizontal shear** is a shear about the x-axis, and a **vertical shear** is a shear about the y-axis.

A shear about a line through the origin is a linear transformation.

Some types of linear transformation

The following transformations are linear transformations.

- The (k,l)-scaling. Its matrix is
$$\begin{pmatrix} k & 0 \\ 0 & l \end{pmatrix}.$$
- The k-dilation. Its matrix is
$$\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}.$$
- The horizontal shear with shear factor k. Its matrix is
$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}.$$
- The vertical shear with shear factor $-k$. Its matrix is
$$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}.$$
- The rotation through the angle θ about the origin. Its matrix is
$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$
- The reflection in the line through the origin with angle of inclination α. Its matrix is
$$\begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}.$$
- A flattening whose image set is the origin or a line through the origin. Its matrix has determinant zero.



To recognise the type of a linear transformation f from its matrix $\mathbf{A}$

Try the following.

- Compare $\mathbf{A}$ with the list on page 91.
- Use one of the criteria on page 90 to check whether f is a flattening.
- Use the determinant.
  - If $\det \mathbf{A} \neq 1$, then f is not a shear or a rotation.
  - If $\det \mathbf{A} \neq -1$, then f is not a reflection.
- If $\det \mathbf{A} = \pm 1$, check whether f is a rotation or a reflection. It is one of these if the columns $\mathbf{i}'$ and $\mathbf{j}'$ of $\mathbf{A}$ are such that $|\mathbf{i}'| = |\mathbf{j}'| = 1$ and $\mathbf{i}' \cdot \mathbf{j}' = 0$ (that is, they are perpendicular unit vectors). If so, then f is
  - a rotation if $\det \mathbf{A} = 1$
  - a reflection if $\det \mathbf{A} = -1$.

To find the image set of a flattening $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$

- If $\mathbf{A} = \mathbf{0}$, then the image set of f contains the origin alone.
- Otherwise, the image set of f is the line that passes through the origin and the points (a, c) and (b, d), where $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Composite and inverse linear transformations

If f and g are linear transformations represented by the matrices $\mathbf{A}$ and $\mathbf{B}$ respectively, then $g \circ f$ is the linear transformation represented by the matrix $\mathbf{BA}$.

If f is an invertible linear transformation represented by the matrix $\mathbf{A}$, then its inverse f^{-1} is the linear transformation represented by the matrix $\mathbf{A}^{-1}$.

To find the equation of the image $f(\mathcal{C})$ of a curve $\mathcal{C}$ under an invertible linear transformation $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$

1. Find, in terms of x and y, the coordinates of the point $f^{-1}(x, y)$.
2. Substitute these coordinates into the equation of $\mathcal{C}$, and simplify.

To use conjugation to find the matrix of the shear with factor k about the line ℓ through the origin, where ℓ is neither horizontal nor vertical

1. Find the matrix $\mathbf{H}$ of a rotation that maps ℓ to the x-axis.
2. Find the matrix $\mathbf{G}$ of the horizontal shear with shear factor k.
3. The required matrix is $\mathbf{H}^{-1}\mathbf{G}\mathbf{H}$.

Affine transformations

An **affine transformation** is a transformation of the form

$$f(\mathbf{x}) = \mathbf{Ax} + \mathbf{a},$$

where $\mathbf{A}$ is a 2×2 matrix and $\mathbf{a}$ is a 2×1 column vector.

To find an affine transformation from the images of $\mathbf{0}$, $\mathbf{i}$ and $\mathbf{j}$

The affine transformation f that maps the points with position vectors $\mathbf{0}$, $\mathbf{i}$ and $\mathbf{j}$ to the points with position vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, respectively, is

$$f(\mathbf{x}) = \mathbf{Ax} + \mathbf{a},$$

where $\mathbf{A}$ is the 2×2 matrix whose first column is $\mathbf{b} - \mathbf{a}$ and whose second column is $\mathbf{c} - \mathbf{a}$.

Here (as usual), $\mathbf{0}$, $\mathbf{i}$ and $\mathbf{j}$ are the position vectors of the points $(0, 0)$, $(1, 0)$ and $(0, 1)$, respectively.

Properties of one-to-one affine transformations

The five listed properties of a one-to-one linear transformation $f(\mathbf{x}) = \mathbf{Ax}$ on page 90 also apply to a one-to-one affine transformation $f(\mathbf{x}) = \mathbf{Ax} + \mathbf{a}$.

Affine transformations that are not one-to-one

The criteria for an affine transformation $f(\mathbf{x}) = \mathbf{Ax} + \mathbf{a}$ to be a flattening are the same as those for a linear transformation $f(\mathbf{x}) = \mathbf{Ax}$ (see page 90).

Composite and inverse affine transformations

A composite of affine transformations is an affine transformation.

If $f(\mathbf{x}) = \mathbf{Ax} + \mathbf{a}$ is an invertible affine transformation, then the inverse of f is the affine transformation $f^{-1}(\mathbf{x}) = \mathbf{A}^{-1}\mathbf{x} - \mathbf{A}^{-1}\mathbf{a}$.

To use conjugation to find an affine transformation

To find the rotation through the angle θ about the point C

1. Find the translation h that maps C to the origin.
2. Find the rotation g through θ about the origin.
3. The required affine transformation is $h^{-1} \circ g \circ h$.

To find the reflection in the line ℓ

1. Find a translation h that maps ℓ to a line through the origin.
2. Find the reflection g in this line through the origin.
3. The required affine transformation is $h^{-1} \circ g \circ h$.

A similar method can be used to find other affine transformations, such as shears about lines that do not pass through the origin.

To determine the type of an affine transformation f that is an isometry

If the matrix in the rule of f is the identity matrix, then f is a translation. Otherwise, solve the equation $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ to find the fixed points of f, and use the number of fixed points to identify what type of isometry f is.

MST125 Unit 7 Topics in calculus

Division of polynomials

The division theorem for polynomial expressions

Suppose that $a(x)$ and $b(x)$ are polynomial expressions and $b(x)$ is not the expression 0. Then there are unique polynomial expressions $q(x)$ and $r(x)$ such that

$$a(x) = q(x)b(x) + r(x),$$

where the degree of $r(x)$ is smaller than the degree of $b(x)$.

The **quotient** of the division of $a(x)$ by $b(x)$ is the expression $q(x)$.

The **remainder** of the division is the expression $r(x)$.

To use polynomial long division to divide a polynomial $a(x)$ by a polynomial $b(x)$

Example:

$a(x) = 3x^3 - 2x^2 + 3$

$b(x) = x + 1$

$$\begin{array}{r|rrrr}
 & 3x^2 & -5x & +5 & \\ \cline{2-5}
x+1 & 3x^3 & -2x^2 & +0x & +3 \\
 & 3x^3 & +3x^2 & & \\ \cline{2-3}
 & & -5x^2 & +0x & +3 \\
 & & -5x^2 & -5x & \\ \cline{3-4}
 & & & 5x & +3 \\
 & & & 5x & +5 \\ \cline{4-5}
 & & & & -2
\end{array}$$

$q(x) = 3x^2 - 5x + 5$

$r(x) = -2$

1. Write down $b(x)\,\overline{\big)\,a(x)}$. Include any 'missing' terms in the polynomial expressions by using the coefficient 0.
2. Find an expression kx^n, where k is a constant and n is a positive integer or 0, such that $kx^n \times b(x)$ has the same highest power term as $a(x)$. Write kx^n at the top, lining it up with the term underneath with the same power.
3. Write $kx^n \times b(x)$ under $a(x)$, lining up terms with the same powers.
4. Subtract the bottom polynomial expression from the one above to obtain a new polynomial expression, writing it at the bottom and lining up terms with the same powers.
5. Repeat steps 2 to 4 with this new polynomial expression instead of $a(x)$ (but still using $b(x)$), to obtain another new polynomial expression. Include the sign (plus or minus) of the term kx^n when you write it at the top.
6. Keep repeating steps 2 to 4 with each new polynomial expression in turn, until the degree of the new polynomial expression obtained by subtracting is smaller than the degree of $b(x)$.
7. The quotient $q(x)$ is the polynomial expression at the top, and the remainder $r(x)$ is the final new polynomial expression at the bottom.

Partial fractions

Rational expressions

A **rational expression** in x is an expression of the form $\dfrac{a(x)}{b(x)}$, where $a(x)$ and $b(x)$ are polynomial expressions, and $b(x)$ is not the expression 0.

It is **proper** if the degree of $a(x)$ is smaller than the degree of $b(x)$, and **improper** otherwise.

Partial fraction expansion of a rational expression

The **partial fraction expansion** of a proper rational expression $f(x)$ is the sum of expressions of the form

$$\frac{A}{(ax+b)^n} \quad \text{and} \quad \frac{Ax+B}{(ax^2+bx+c)^n}$$

that is equal to $f(x)$. Here n is a positive integer, a, b, c, A and B are constants, and the discriminant $b^2 - 4ac$ of each quadratic expression $ax^2 + bx + c$ is negative.

The **partial fraction expansion** of an improper rational expression $f(x)$ is the sum of a polynomial expression and expressions of the form above that is equal to $f(x)$.

The **partial fractions** of a rational expression $f(x)$ are the terms in its partial fraction expansion.

Every rational expression has a partial fraction expansion.

To find the partial fraction expansion of a rational expression $\dfrac{a(x)}{b(x)}$ without repeated quadratic factors in the denominator $b(x)$

- If the rational expression is proper, then use the method on the following page.
- If the rational expression is improper, then proceed as follows.
  1. Use polynomial long division to write
     $$a(x) = q(x)b(x) + r(x).$$
  2. Divide through by $b(x)$ to obtain
     $$\frac{a(x)}{b(x)} = q(x) + \frac{r(x)}{b(x)},$$
     where the degree of $r(x)$ is smaller than the degree of $b(x)$.
  3. Use the method on the following page to find the partial fraction expansion of the proper rational expression $\dfrac{r(x)}{b(x)}$.

To integrate a rational expression

Find its partial fraction expansion and integrate each term separately. (Sometimes an alternative method may work too.)

To find the partial fraction expansion of a proper rational expression without repeated quadratic factors in the denominator

1. Write down the form of the partial fraction expansion, using a different letter to denote each unknown constant.
   - Each linear factor $ax+b$ that is not repeated gives rise to a partial fraction of the form $\dfrac{A}{ax+b}$, where A is a constant.
   - Each repeated linear factor $(ax+b)^n$ gives rise to n partial fractions of the form
   $$\frac{A_1}{ax+b}+\frac{A_2}{(ax+b)^2}+\cdots+\frac{A_n}{(ax+b)^n},$$
   where $A_1, A_2, \ldots, A_n$ are constants.
   - Each quadratic factor ax^2+bx+c that is not repeated and whose discriminant b^2-4ac is negative gives rise to a partial fraction of the form $\dfrac{Ax+B}{ax^2+bx+c}$, where A and B are constants.
2. Find as many of the constants as you can by using the cover-up method.
3. - If only one constant remains unknown, then find it by using the method of substituting values.
   - If more than one constant remains unknown, then find them by using the method of equating coefficients.

Methods for determining the constants in partial fractions

Cover-up method Choose a factor of the form $(ax+b)^n$, where $n \geq 1$, of the denominator of the original rational expression (such that no larger power of $ax+b$ is also a factor). Cover up this factor, and substitute into the uncovered part the value of x that makes the covered-up factor equal to 0. The resulting number is the constant A in the partial fraction $A/(ax+b)^n$.

Preliminary step for the next two methods

Multiply the equation

$$\text{original rational expression} = \begin{array}{l}\text{partial fraction expansion}\\ \text{including unknown constants}\end{array}$$

through by the denominator of the original rational expression to obtain an equation of the form

$$\text{polynomial expression} = \text{polynomial expression}.$$

Method of substituting values Substitute into the equation a value of x that eliminates all the unknown constants except one. Solve the resulting equation.

Method of equating coefficients Use the fact that on both sides of the equation the constant term is the same, as is the coefficient of x, as is the coefficient of x^2, and so on, to obtain equations in the unknown constants. Solve these equations.

Features of graphs of rational functions

A **rational function** is a function whose rule is of the form $f(x) = \dfrac{a(x)}{b(x)}$, where $a(x)$ and $b(x)$ are polynomial expressions and $b(x)$ is not the expression 0.

The graph of a rational function consists of one or more continuous pieces.

Domain

Unless otherwise stated, the domain of a rational function $f(x) = \dfrac{a(x)}{b(x)}$, where $a(x)$ and $b(x)$ are polynomial expressions, consists of all real numbers except the values of x for which $b(x) = 0$.

Intercepts

The x- and y-**intercepts** of a graph are the values of x and y where the graph crosses the x-axis and the y-axis, respectively.

Stationary points and intervals on which a function is increasing or decreasing

You can find the stationary points of a rational function f, their natures, and the intervals on which f is increasing or decreasing by constructing a table of signs for $f'(x)$, see page 49. For how to construct a table of signs, see page 36.

Asymptotic behaviour

The **asymptotic behaviour** of a graph is its behaviour as you trace your pen tip further and further along each infinitely long piece of the graph.

An **asymptote** of a graph is a line that the graph approaches arbitrarily closely as you trace your pen tip along it further and further away from the origin.

A **horizontal**, **vertical** or **slant asymptote** is an asymptote that is a horizontal, vertical or slant line, respectively.

Asymptotic behaviour of a rational function as $x \to \infty$ and as $x \to -\infty$

The graph of a rational function f has one of the following five types of behaviour as $x \to \infty$ and as $x \to -\infty$:

- $f(x) \to a$ as $x \to \infty$ and as $x \to -\infty$, for some constant number a
- $f(x) \to \infty$ as $x \to \infty$ and as $x \to -\infty$
- $f(x) \to -\infty$ as $x \to \infty$ and as $x \to -\infty$
- $f(x) \to \infty$ as $x \to \infty$ and $f(x) \to -\infty$ as $x \to -\infty$
- $f(x) \to -\infty$ as $x \to \infty$ and $f(x) \to \infty$ as $x \to -\infty$

The notation here has the following meanings:

$\to \infty$ means 'becomes arbitrarily large in magnitude and positive'
$\to -\infty$ means 'becomes arbitrarily large in magnitude and negative'
$\to a$ means 'approaches a'.

Even and odd functions

A function f is **even** if $f(-x) = f(x)$ for all x in the domain of f.
A function f is **odd** if $f(-x) = -f(x)$ for all x in the domain of f.

To determine the vertical asymptotes of a rational function f

Each number a for which the denominator of the rational expression $f(x)$ is zero gives a vertical asymptote $x = a$.

To determine the asymptotic behaviour of a rational function f as $x \to \infty$ and as $x \to -\infty$, and find any horizontal or slant asymptote

1. Consider the function g such that $g(x)$ is obtained by deleting all the terms except the **dominant term** (the term with the highest power of x) in each of the numerator and the denominator of the rational expression $f(x)$. Determine which of the five types of asymptotic behaviour listed on the previous page apply to g (including the value of the constant a, if applicable). The same behaviour applies to f.
2. If $f(x) \to a$ as $x \to \infty$ and as $x \to -\infty$ (this happens only when the degree of the numerator of the rational expression $f(x)$ is less than or equal to the degree of the denominator), then the line $y = a$ is a horizontal asymptote.
3. If the degree of the numerator of the rational expression $f(x)$ is one more than the degree of the denominator, then the graph of f has a slant asymptote (otherwise it has no slant asymptote). To determine the equation of the slant asymptote, use polynomial long division to write $f(x) = q(x) + h(x)$, where $q(x)$ is a polynomial expression and $h(x)$ is a proper rational expression. The slant asymptote has equation $y = q(x)$. To determine how the graph approaches $y = q(x)$, consider the sign of $h(x)$ as $x \to \infty$ and as $x \to -\infty$.

To sketch the graph of a rational function

1. Find the domain and then draw and label any vertical asymptotes.
2. Find and label the x- and y-intercepts, if any.
3. Find and label any stationary points, indicating their natures by drawing small curves through them, and find the intervals on which the function is increasing or decreasing.
4. Sketch the shape of the graph near the vertical asymptotes, if any.
5. Work out the asymptotic behaviour as $x \to \infty$ and as $x \to -\infty$. Draw and label any horizontal or slant asymptote, and sketch the shape of the graph near the far left and far right of any such asymptote.
6. Check whether the function is even or odd, or neither.
7. In addition to these steps you can also
   - plot particular points on the graph
   - work out the intervals on which the function is positive or negative
   - use the second derivative to find the intervals on which the function is concave up or concave down
   - use the second derivative test to check the nature of some stationary points.

Hyperbolic functions

sinh, cosh and tanh

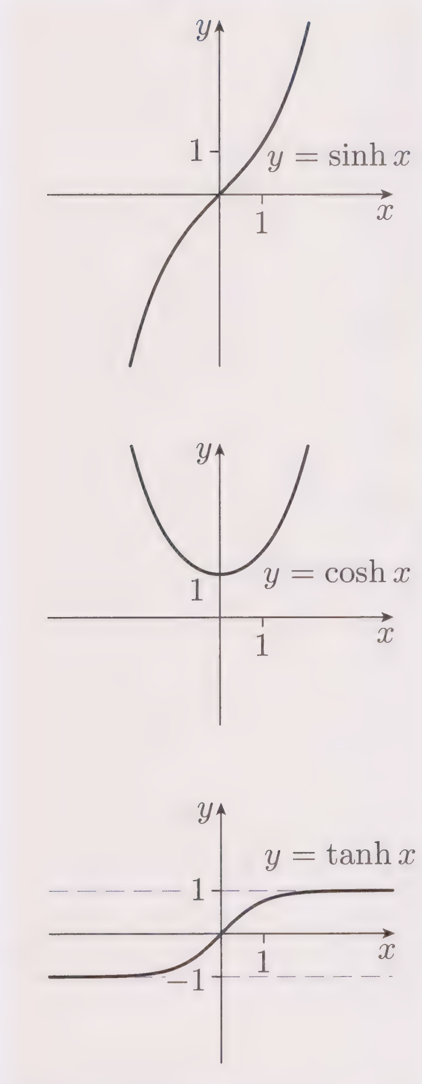
The **hyperbolic functions sinh**, **cosh** and **tanh** are given by

$$\sinh x = \tfrac{1}{2}(e^x - e^{-x})$$

$$\cosh x = \tfrac{1}{2}(e^x + e^{-x})$$

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

sinh and tanh are odd functions and cosh is an even function.



Inverse hyperbolic functions

The **inverse sinh function** $\sinh^{-1}$ has domain $\mathbb{R}$ and rule

$$\sinh^{-1} x = y,$$

where y is the real number such that $\sinh y = x$.

The **inverse cosh function** $\cosh^{-1}$ has domain $[1, \infty)$ and rule

$$\cosh^{-1} x = y,$$

where y is the number in the interval $[0, \infty)$ such that $\cosh y = x$.

The **inverse tanh function** $\tanh^{-1}$ has domain $(-1, 1)$ and rule

$$\tanh^{-1} x = y,$$

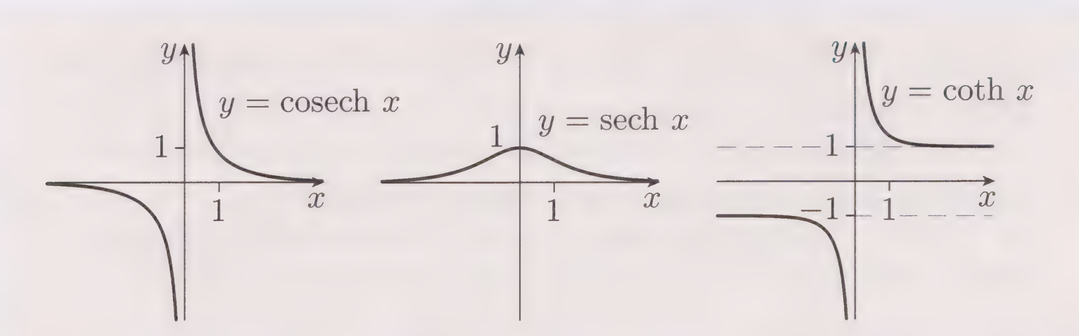
where y is the real number such that $\tanh y = x$.

For the graphs of the inverse hyperbolic functions, see page 11.

cosech, sech and coth

$$\operatorname{cosech} x = \frac{1}{\sinh x} \qquad \operatorname{sech} x = \frac{1}{\cosh x} \qquad \coth x = \frac{1}{\tanh x}$$

($\operatorname{cosech} x$ and $\coth x$ are not defined when $x = 0$.)



For hyperbolic identities, see page 6. For the derivatives and the indefinite integrals of hyperbolic functions, see page 8.

Integrating trigonometric expressions

To find $\int \sin(mx)\cos(nx)\,dx$, where m and n are integers

Use a product to sum identity (page 5).

To find $\int \sin^m x \cos^n x\,dx$, where m and n are non-negative integers

- If the power of $\sin x$ is odd, then separate out one factor $\sin x$ and use the identity $\sin^2 x + \cos^2 x = 1$ to express the remaining power of $\sin x$ in terms of $\cos x$. Then use integration by substitution.
- If the power of $\cos x$ is odd, then separate out one factor $\cos x$ and use the identity $\sin^2 x + \cos^2 x = 1$ to express the remaining power of $\cos x$ in terms of $\sin x$. Then use integration by substitution.
- If the powers of $\sin x$ and $\cos x$ are both even, then use the half-angle identities (page 5). The identity $\sin x \cos x = \frac{1}{2}\sin(2x)$ may also help.

Trigonometric and hyperbolic substitutions

Trigonometric substitutions

Expression	Substitution	Identity
$a^2 - x^2$	$x = a\sin u$	$1 - \sin^2 u = \cos^2 u$
$x^2 - a^2$	$x = a\sec u$	$\sec^2 u - 1 = \tan^2 u$
$a^2 + x^2$	$x = a\tan u$	$1 + \tan^2 u = \sec^2 u$

Hyperbolic substitutions

Expression	Substitution	Identity
$x^2 - a^2$	$x = a\cosh u$	$\cosh^2 u - 1 = \sinh^2 u$
$a^2 + x^2$	$x = a\sinh u$	$1 + \sinh^2 u = \cosh^2 u$

Use the identity once you have made the substitution in the integrand.

Choosing a method for finding an integral

Try the list of suggestions on page 57, and the further suggestions below.

- Is the integrand a rational expression? If so, then find its partial fraction expansion, and integrate the partial fractions separately.
- Does the integrand contain one of the expressions $a^2 - x^2$, $x^2 - a^2$ or $a^2 + x^2$; in particular, does it contain the square root of one of these expressions? If so, then try a trigonometric substitution.
- Does a trigonometric substitution still leave you with a difficult integral? If so, and the integrand contains one of the expressions $a^2 - x^2$ or $a^2 + x^2$ (in particular, if it contains the square root of one of these expressions), then try a hyperbolic substitution instead of a trigonometric substitution.
- Does the integrand contain hyperbolic functions? If so, try methods similar to those for trigonometric functions, such as using identities.

MST125 Unit 8 Differential equations

Terminology for differential equations

A **differential equation** is an equation that involves an unknown function and its derivatives.

The **order** of a differential equation is the order of the highest derivative appearing in the equation. (A derivative of **order** n is an nth derivative.)

A **first-order** differential equation is a differential equation of order 1.

A **solution** of a differential equation is a function $y = y(x)$ that satisfies the differential equation.

The **general solution** of a differential equation is a solution that contains one or more arbitrary constants, such that any solution of the differential equation is obtained by choosing particular values for the arbitrary constants.

A **particular solution** of a differential equation is obtained by choosing a particular value for each arbitrary constant in the general solution.

An **explicit solution** of a differential equation is a solution in the form of an equation whose subject is the dependent variable.

An **implicit solution** of a differential equation is a solution in the form of an equation relating the dependent and independent variables that does not have the dependent variable as a subject.

An **initial condition** for a differential equation is a requirement that when the independent variable takes some specified value, the dependent variable takes some specified value.

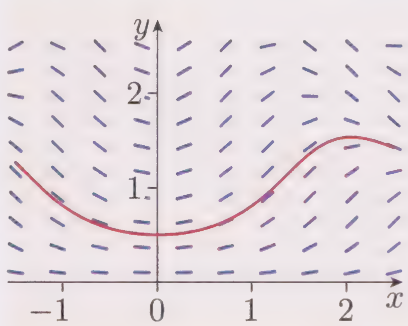
An **initial value problem** involves finding the particular solution of a differential equation that satisfies a given initial condition.

Solving a differential equation means finding its general solution.

Solving an initial value problem means finding the particular solution of the given differential equation that satisfies the given initial condition.

A **parameter** in a differential equation is a constant represented by a symbol. Each value of the parameter gives a specific differential equation.

A **direction field** for a differential equation of the form $\mathrm{d}y/\mathrm{d}x = f(x, y)$ is a set of many short line segments plotted in the x, y-plane such that the gradient of each line segment is $f(x, y)$ where (x, y) is the centre of the line segment. The line segments provide a guide to the shape of the graph of any particular solution of the differential equation.



A **numerical solution** of a differential equation in the variables x and y is a sequence of points in the x, y-plane that when joined together in sequence form an approximation to the graph of a particular solution of the differential equation.

An **analytic solution** of a differential equation is a general solution or a particular solution given in implicit or explicit form.

Solving differential equations

Directly integrable differential equations

A **directly integrable** first-order differential equation is one of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x).$$

To solve it, integrate both sides with respect to x.

Separable differential equations

A **separable** first-order differential equation is one of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y).$$

To solve it, use **separation of variables**:

1. Separate the variables and put integral signs on both sides to obtain
$$\int \frac{1}{g(y)}\,\mathrm{d}y = \int f(x)\,\mathrm{d}x.$$
2. Carry out the two integrations, introducing one arbitrary constant, to obtain the general solution in implicit form.
3. If possible, manipulate the resulting equation to make y the subject, thus expressing the general solution in explicit form.

Linear differential equations

A **linear** first-order differential equation is one of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} + g(x)y = h(x).$$

It is **homogeneous** if $h(x) = 0$ for all x, and **inhomogeneous** or **non-homogeneous** otherwise.

To solve it, use an integrating factor:

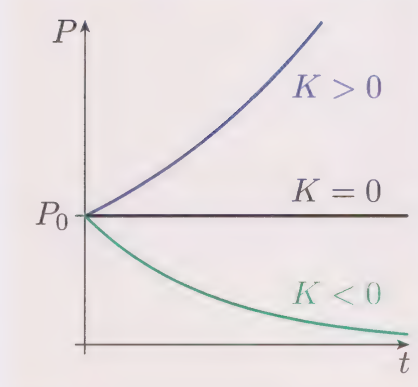
1. Write down the **integrating factor**
$$p(x) = \exp\left(\int g(x)\,\mathrm{d}x\right).$$
Find the integral, if possible, but do not include a constant of integration.
2. Write down the general solution, given by the formula
$$y = \frac{1}{p(x)}\left(\int p(x)h(x)\,\mathrm{d}x\right).$$
Find the integral, if possible, and include the constant of integration inside the brackets.

To solve an initial value problem

1. Find the general solution of the differential equation.
2. Use the initial condition to find the value(s) of the arbitrary constant(s).

Applications of differential equations

Exponential model for population change



Let P be the size of a population at time t. It may be appropriate to model the change in P by the differential equation and initial condition

$$\frac{dP}{dt} = KP, \quad \text{where } P = P_0 \text{ when } t = 0.$$

Here K is a constant, called the **proportionate growth rate** for the population size.

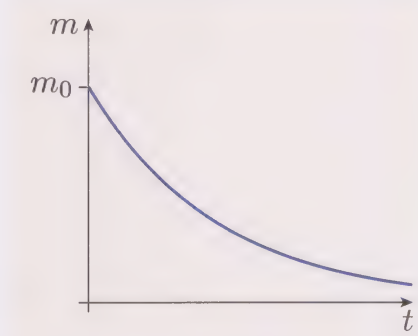
With this model, the population size P at time t is given by

$$P = P_0 e^{Kt}.$$

If $K > 0$, then the population size increases, with doubling time $\dfrac{\ln 2}{K}$.

If $K < 0$, then the population size decreases, with halving time $-\dfrac{\ln 2}{K}$.

Exponential model for radioactive decay



Let m be the mass of a radioactive substance at time t. Then the change in m can be modelled by the differential equation and initial condition

$$\frac{dm}{dt} = -km, \quad \text{where } m = m_0 \text{ when } t = 0.$$

Here k is a positive constant, called the **decay constant**.

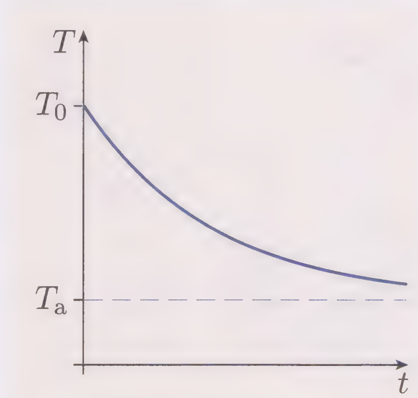
With this model, the mass m of the substance at time t is given by

$$m = m_0 e^{-kt}.$$

The **half-life** of a radioactive substance is the time that it takes for the mass of the substance to halve.

If a radioactive substance decays with decay constant k, then its half-life T is given by $T = \dfrac{\ln 2}{k}$.

Newton's law of cooling



Let T be the temperature of a cooling body at time t, and let T_a be the temperature of the surrounding medium (the **ambient temperature**). According to Newton's law of cooling, the change in T is modelled by the differential equation and initial condition

$$\frac{dT}{dt} = -\lambda(T - T_a), \quad \text{where } T = T_0 \text{ when } t = 0.$$

Here λ is a positive constant that depends on the body and the surrounding medium.

With this model, if the ambient temperature T_a is constant then the temperature T of the body at time t is given by

$$T = T_a + (T_0 - T_a)e^{-\lambda t}.$$

MST125 Unit 9 Mathematical language and proof

Statements

A **statement**, or **proposition**, is an assertion that is either true or false.

A **proof** is a logical argument establishing that a statement is true.

A **conjecture** is a statement that seems likely to be true but has not been proved.

A **theorem** is a true mathematical statement of some importance.

A **lemma** is a true statement that is not important enough to be called a theorem. It is usually a statement that can be used to prove that a more important statement is true.

A **corollary** is a mathematical statement whose truth can be established relatively easily from a related mathematical statement (often a theorem) that is known to be true.

A **result** is a mathematical statement that has been proved (or, more generally, the outcome of some mathematical process).

A **variable proposition** is a mathematical statement that is either true or false depending on the values of the variables in the statement.

An **existential statement** is a statement asserting that a variable proposition is true for at least one value of each of its variables. It includes the phrase 'there exists' or another form of words with the same meaning.

A **universal statement** is a statement asserting that a variable proposition is true for every value of each of its variables. It includes the phrase 'for all' or another form of words with the same meaning.

An **existential quantifier** is the symbol $\exists$, or any words or phrases with the same meaning, such as 'there exists' and 'there are'.

A **universal quantifier** is the symbol $\forall$, or any words or phrases with the same meaning, such as 'for all' and 'for every'.

Combining statements

Truth tables

A **truth table** is a table that sets out how the truth of one or more statements depends on the truth of one or more other statements.

P	Q	P AND Q
true	true	true
true	false	false
false	true	false
false	false	false

AND and OR

The statement P AND Q is true if both P and Q are true, and false otherwise.

The statement P OR Q is true if at least one of P or Q is true, and false otherwise.

P	Q	P OR Q
true	true	true
true	false	true
false	true	true
false	false	false

Negating statements

The statement NOT P (the **negation** of P) is false if P is true and true if P is false.

P	NOT P
true	false
false	true

Negating combinations of statements

The negation of P AND Q is (NOT P) OR (NOT Q).

The negation of P OR Q is (NOT P) AND (NOT Q).

Negating existential and universal statements

1. Universal quantifiers become existential quantifiers, and vice versa.
2. The underlying statement to which a quantifier applies is negated.

This is illustrated by the following forms of statement and their negations.

- statement: $P(x)$ is true for all $x \in A$
  negation: there exists $x \in A$ for which $P(x)$ is false
- statement: there exists $x \in A$ for which $P(x)$ is true
  negation: $P(x)$ is false for all $x \in A$

Implications and equivalences

An **implication** is a statement of the form 'if P, then Q', where P and Q are statements. Other ways of writing this implication include: '$P \Rightarrow Q$', 'P implies Q', 'Q whenever P', 'Q follows from P', 'P is sufficient for Q' and 'Q is necessary for P'.

P	Q	if P, then Q
true	true	true
true	false	false
false	true	true
false	false	true

The **converse** of an implication 'if P, then Q' is the implication 'if Q, then P'. If an implication is true, then its converse may or may not be true.

The **contrapositive** of an implication 'if P, then Q', is the equivalent implication 'if NOT Q, then NOT P'.

An **equivalence** is a statement of the form 'P if and only if Q'. It asserts that both the implication 'if P, then Q' and its converse 'if Q, then P' are true. Other ways of writing this equivalence include: '$P \Leftrightarrow Q$', 'P iff Q', 'P is equivalent to Q' and 'P is necessary and sufficient for Q'.

Convention for implications and equivalences

In a universal statement whose underlying variable proposition is an implication or equivalence, the quantifier part is often omitted. Thus:

- the universal statement 'for all $x \in X$, if $P(x)$ then $Q(x)$' is usually written simply as 'if $P(x)$, then $Q(x)$';
- the universal statement 'for all $x \in X$, $P(x)$ if and only if $Q(x)$' is usually written simply as '$P(x)$ if and only if $Q(x)$'.

Characterisation of odd and even integers

- An integer n is **odd** if and only if $n = 2k + 1$ for some integer k.
- An integer n is **even** if and only if $n = 2k$ for some integer k.

Direct proofs

Key deductive step

If the statements P and $P \Rightarrow Q$ are both true, then the statement Q is also true.

Deductive step using equivalence

If the statements Q and $P \Leftrightarrow Q$ are both true, then the statement P is also true.

(These deductive steps may use additional well-known results or definitions without comment: for example, results about manipulating equations or inequalities are often used in this way.)

To prove an implication $P \Rightarrow Q$ using a sequence of deductions

1. Assume that P is true.
2. Identify a statement P_1 for which you know that $P \Rightarrow P_1$ is true, and deduce that P_1 is true.
3. Identify a statement P_2 for which you know that $P_1 \Rightarrow P_2$ is true, and deduce that P_2 is true.
4. Continue this process until you obtain a statement P_n for which you know that $P_n \Rightarrow Q$ is true.
5. Deduce that Q is true.

To prove an equivalence $P \Leftrightarrow Q$

EITHER:

Prove that each of

$$P \Rightarrow Q \quad \text{and} \quad Q \Rightarrow P$$

is true.

OR:

1. Write down P.
2. Identify a statement P_1 for which you know that $P \Leftrightarrow P_1$ is true.
3. Identify a statement P_2 for which you know that $P_1 \Leftrightarrow P_2$ is true.
4. Continue this process until you obtain a statement P_n for which you know that $P_n \Leftrightarrow Q$ is true.
5. Deduce that $P \Leftrightarrow Q$ is true.

To prove a statement P using a sequence of equivalences

1. Write down P.
2. Identify a statement P_1 for which you know that $P \Leftrightarrow P_1$ is true.
3. Identify a statement P_2 for which you know that $P_1 \Leftrightarrow P_2$ is true.
4. Continue this process until you obtain a statement Q for which you know that $P_n \Leftrightarrow Q$ is true and Q is true.
5. Deduce that P is true.

Proof by induction

To prove by mathematical induction that a variable proposition $P(n)$ is true for all $n \in \mathbb{N}$

1. Show that $P(1)$ is true.
2. Show that the implication $P(k) \Rightarrow P(k+1)$ is true for $k = 1, 2, \ldots$.
3. Deduce that $P(n)$ is true for $n \in \mathbb{N}$.

To prove by mathematical induction that a variable proposition $P(n)$ is true for all integers $n \geq N$

1. Show that $P(N)$ is true.
2. Show that the implication $P(k) \Rightarrow P(k+1)$ is true for $k \geq N$.
3. Deduce that $P(n)$ is true for $n \geq N$.

The **inductive** step in mathematical induction is step 2 in each list above.

Indirect proofs

To prove a statement P using proof by contradiction

1. Assume that P is false. In other words, assume that the negation Q of P is true.
2. Deduce a **contradiction** (a conclusion that is known to be false) by constructing a sequence of statements $P_1, \ldots, P_n$ such that all the implications $Q \Rightarrow P_1,\ P_1 \Rightarrow P_2, \ldots,\ P_{n-1} \Rightarrow P_n$ are true, but P_n is false.
3. Deduce that Q is false and hence that P is true.

To prove an implication 'if P, then Q' using proof by contraposition

1. Write down the contrapositive implication

   if NOT Q, then NOT P.
2. Use a direct method of proof to show that the contrapositive is true.
3. Deduce that the original implication is true.

Proofs by construction and exhaustion

To prove by construction that an existential statement is true

Find (construct) a value of each variable such that the underlying variable proposition is true.

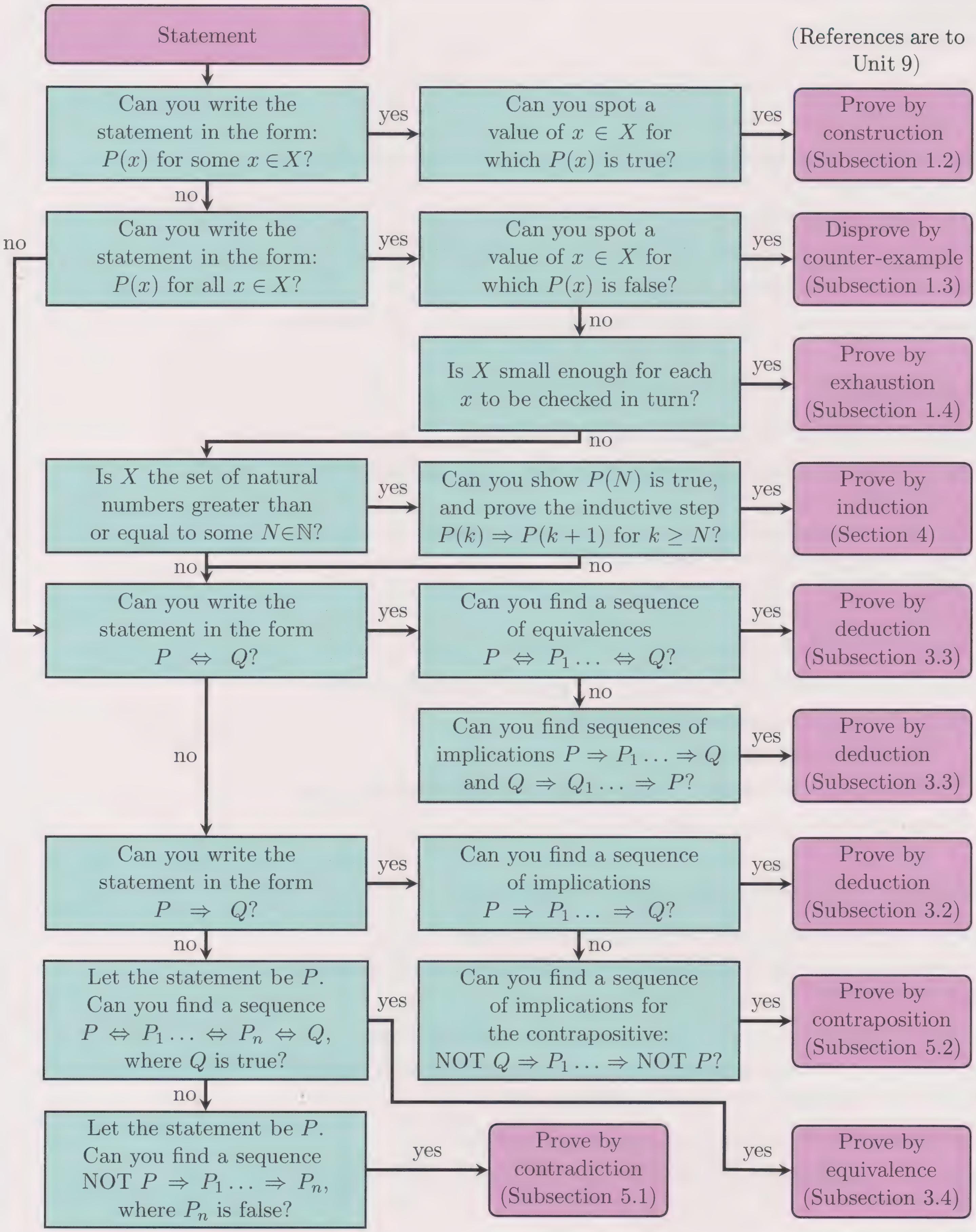
To prove by counter-example that a universal statement is false

Find (construct) a **counter-example**; that is, one value of each variable such that the underlying variable proposition is false.

To prove by exhaustion that a universal statement is true

Check that the underlying variable proposition is true for each value of the variables in turn.

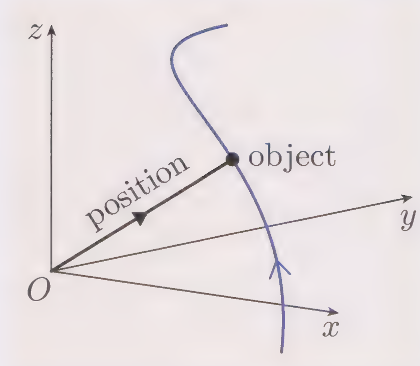
Deciding which method of proof to use



MST125 Unit 10 Dynamics

Position, velocity and acceleration

The **position** of an object is its position vector.



The **velocity** of an object is the rate of change of its position with respect to time.

The **acceleration** of an object is the rate of change of its velocity with respect to time.

Distance is the magnitude of position.
Speed is the magnitude of velocity.
The magnitude of acceleration is also called acceleration.

An **equation of motion** is an equation relating two or more of the four quantities time, position, velocity and acceleration.

Kinematics is the study of motion without consideration of the causes of the motion. It involves using equations of motion.

Position, velocity and acceleration in one dimension

For motion along a straight line, we represent the vector quantities position, velocity and acceleration by scalars, with direction indicated by the signs of the scalars.

If a particle is moving along a straight line, with position x, velocity v and acceleration a at time t, then the following equations hold.

$$v = \frac{\mathrm{d}x}{\mathrm{d}t} \qquad a = \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\mathrm{d}^2x}{\mathrm{d}t^2} \qquad x = \int v\,\mathrm{d}t \qquad v = \int a\,\mathrm{d}t \qquad a = v\frac{\mathrm{d}v}{\mathrm{d}x}$$

- If the velocity and acceleration of an object are both positive or both negative, then the speed of the object is increasing.
- If the velocity and acceleration of an object have opposite signs, then the speed of the object is decreasing.

Graphs of position, velocity and acceleration against time

A **position–time graph** is a graph of position against time.

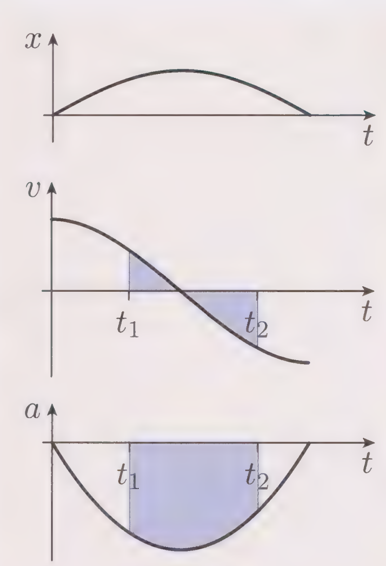
- The gradient of a position–time graph is the velocity of the object.

A **velocity–time graph** is a graph of velocity against time.

- The gradient of a velocity–time graph is the acceleration of the object.
- The signed area between a velocity–time graph and the time axis from time t_1 to time t_2 is the change in the position of the object from time t_1 to time t_2.

An **acceleration–time graph** is a graph of acceleration against time.

- The signed area between an acceleration–time graph and the time axis from time t_1 to time t_2 is the change in the velocity of the object from time t_1 to time t_2.



Equations of motion for constant acceleration along a straight line

If a particle is moving along a straight line with constant acceleration a, with position x and velocity v at time t, and with $x = 0$ and $v = v_0$ at time $t = 0$, then the following equations hold.

$$x = v_0 t + \tfrac{1}{2}at^2 \qquad x = \tfrac{1}{2}(v_0 + v)t \qquad x = vt - \tfrac{1}{2}at^2$$

$$v = v_0 + at \qquad v^2 = v_0^2 + 2ax$$

Piecewise constant acceleration is acceleration that is constant for a period of time, then constant again (but with a possibly different value) for a subsequent period of time, and so on.

Newton's second law of motion

II If a particle of constant mass m is acted on by a resultant force $\mathbf{F}$, then its acceleration $\mathbf{a}$ is given by

$$\mathbf{F} = m\mathbf{a}.$$

To solve a dynamics problem involving several forces

1. Make any necessary modelling assumptions.
2. Draw a diagram of the physical situation, annotating it with any relevant information.
3. Choose the positions of the coordinate axes and draw them on the diagram.

   If the motion is along a straight line, then choose one coordinate axis (usually the x-axis) to lie along the line of motion. Usually, choose the origin to be at the initial location of the moving object, if possible.
4. Identify all the forces acting and draw a separate force diagram, stating any known magnitudes of forces. Mark the directions of $\mathbf{i}$ and $\mathbf{j}$ on this diagram.
5. Express the forces in component form, in terms of unknown quantities if necessary, and find the resultant force.
6. Express the acceleration in component form, in terms of unknown components.
7. Apply Newton's second law.
8. Resolve the resulting vector equation to obtain scalar equations.
9. Solve these equations to find the acceleration.
10. Hence find any other required details about the motion.
11. State a conclusion.

Sliding friction

A **sliding friction force** from a flat surface on an object moving across the surface is a force that acts on the object, in the direction parallel to the surface and opposite to the motion. Its magnitude is

$$\mu|\mathbf{N}|,$$

where μ is the **coefficient of sliding friction** (a constant that depends on the two materials in contact), and $\mathbf{N}$ is the normal reaction of the surface on the object.

Vector and scalar functions

A **vector function** is a function whose output values are vectors.
A **scalar function** is a function whose output values are scalars.

Determining position, velocity and acceleration

Determining velocity and acceleration from position

A particle with position $\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$

has velocity $\mathbf{v} = \dfrac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \dfrac{\mathrm{d}x}{\mathrm{d}t}\,\mathbf{i} + \dfrac{\mathrm{d}y}{\mathrm{d}t}\,\mathbf{j} + \dfrac{\mathrm{d}z}{\mathrm{d}t}\,\mathbf{k}$

and acceleration $\mathbf{a} = \dfrac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = \dfrac{\mathrm{d}^2x}{\mathrm{d}t^2}\,\mathbf{i} + \dfrac{\mathrm{d}^2y}{\mathrm{d}t^2}\,\mathbf{j} + \dfrac{\mathrm{d}^2z}{\mathrm{d}t^2}\,\mathbf{k}.$

Determining acceleration from velocity

A particle with velocity $\mathbf{v} = v_x\,\mathbf{i} + v_y\,\mathbf{j} + v_z\,\mathbf{k}$

has acceleration $\mathbf{a} = \dfrac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \dfrac{\mathrm{d}v_x}{\mathrm{d}t}\,\mathbf{i} + \dfrac{\mathrm{d}v_y}{\mathrm{d}t}\,\mathbf{j} + \dfrac{\mathrm{d}v_z}{\mathrm{d}t}\,\mathbf{k}.$

Determining position from velocity

A particle with velocity $\mathbf{v} = v_x\,\mathbf{i} + v_y\,\mathbf{j} + v_z\,\mathbf{k}$

has position $\mathbf{r} = \displaystyle\int \mathbf{v}\,\mathrm{d}t = \left(\int v_x\,\mathrm{d}t\right)\mathbf{i} + \left(\int v_y\,\mathrm{d}t\right)\mathbf{j} + \left(\int v_z\,\mathrm{d}t\right)\mathbf{k}.$

Determining velocity from acceleration

A particle with acceleration $\mathbf{a} = a_x\,\mathbf{i} + a_y\,\mathbf{j} + a_z\,\mathbf{k}$

has velocity $\mathbf{v} = \displaystyle\int \mathbf{a}\,\mathrm{d}t = \left(\int a_x\,\mathrm{d}t\right)\mathbf{i} + \left(\int a_y\,\mathrm{d}t\right)\mathbf{j} + \left(\int a_z\,\mathrm{d}t\right)\mathbf{k}.$

Projectiles

A **projectile** is an object that is propelled through space by a force that ceases after launch.

The **trajectory** of a projectile is its path.

The **range** of a projectile is the distance along the ground that it travels.

The acceleration of a projectile is $\mathbf{a} = -g\,\mathbf{j}$, where g is the magnitude of the acceleration due to gravity, and $\mathbf{j}$ is the Cartesian unit vector that points vertically upwards.

MST125 Unit 11 Eigenvalues and eigenvectors

Terminology for eigenvalues and eigenvectors

An **invariant line** of a linear transformation is a line whose image under the transformation is either the original line, or a point on the original line.

An **eigenvalue** of a square matrix $\mathbf{A}$ is a number λ such that $\mathbf{Ax} = \lambda\mathbf{x}$ for some non-zero vector $\mathbf{x}$. This vector $\mathbf{x}$ is called an **eigenvector** of $\mathbf{A}$ **corresponding** to λ. Any non-zero scalar multiple of $\mathbf{x}$ is also an eigenvector of $\mathbf{A}$ corresponding to λ.

The **eigenvector equation** for the eigenvalue λ of a square matrix $\mathbf{A}$ is the equation $\mathbf{Ax} = \lambda\mathbf{x}$, or equivalently $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$.

The **trace** of a 2×2 matrix $\mathbf{A}$, denoted by $\operatorname{tr}\mathbf{A}$, is the sum of the elements on the **leading diagonal** (top left to bottom right) of $\mathbf{A}$.

The **characteristic equation** of a 2×2 matrix $\mathbf{A}$ is the quadratic equation in λ given by $\lambda^2 - (\operatorname{tr}\mathbf{A})\lambda + \det\mathbf{A} = 0$. The eigenvalues of $\mathbf{A}$ are the solutions of this equation, so $\mathbf{A}$ has 1 or 2 (not always real) eigenvalues.

A **repeated eigenvalue** of a 2×2 matrix $\mathbf{A}$ is an eigenvalue of $\mathbf{A}$ that is a repeated solution of the characteristic equation of $\mathbf{A}$.

Finding eigenvalues and eigenvectors

To find the eigenvalues of a 2×2 matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

1. Find $\operatorname{tr}\mathbf{A} = a + d$ and $\det\mathbf{A} = ad - bc$.
2. Write down the characteristic equation $\lambda^2 - (\operatorname{tr}\mathbf{A})\lambda + \det\mathbf{A} = 0$.
3. Solve the characteristic equation to obtain the eigenvalues of $\mathbf{A}$.

Useful check: The sum of the eigenvalues of $\mathbf{A}$ is equal to $\operatorname{tr}\mathbf{A}$. (A single, repeated eigenvalue must be counted twice in the sum.)

To find an eigenvector of a 2×2 matrix $\mathbf{A}$ corresponding to a given eigenvalue λ of $\mathbf{A}$

1. Write down the eigenvector equation $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, where $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ is an eigenvector to be determined.
2. Rewrite this equation as a pair of simultaneous equations in x and y.
3. These simultaneous equations are together equivalent to a single equation in x and y. Any solution of this equation other than $x = y = 0$ gives an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ.

Eigenvectors of a 2×2 matrix $\mathbf{A}$ corresponding to a real eigenvalue λ

- If $\mathbf{A} \neq \lambda\mathbf{I}$, then the eigenvectors of $\mathbf{A}$ corresponding to λ are the position vectors of the points on a single line through the origin, other than the position vector of the origin itself.
- If $\mathbf{A} = \lambda\mathbf{I}$, then every non-zero vector is an eigenvector of $\mathbf{A}$ corresponding to λ.

Eigenvalues and eigenvectors of special types of matrices

Triangular matrices

A 2×2 **upper triangular matrix** is a matrix of the form $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$.
It has eigenvalues a and d, and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is an eigenvector corresponding to a.

A 2×2 **lower triangular matrix** is a matrix of the form $\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$.
It has eigenvalues a and d, and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is an eigenvector corresponding to d.

A 2×2 **triangular matrix** is a 2×2 upper or lower triangular matrix.

Matrices of some geometric transformations

Transformation	Matrix	Eigenvalues	Eigenvectors
horizontal shear ($k \neq 0$)	$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$	1 (repeated)	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
vertical shear ($k \neq 0$)	$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$	1 (repeated)	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
rotation about O ($\theta \neq n\pi$, $n \in \mathbb{Z}$)†	$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$	no real eigenvalues	no real eigenvectors
reflection in a line through O	$\begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}$	-1, 1	$\begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$, $\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$
k-dilation	$k\mathbf{I}$	k (repeated)	all non-zero vectors
(k,l)-scaling ($k \neq l$)	$\begin{pmatrix} k & 0 \\ 0 & l \end{pmatrix}$	k, l	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

† A rotation about the origin through $n\pi$ is the dilation represented by $\mathbf{I}$ if n is even, and is the dilation represented by $-\mathbf{I}$ if n is odd.

Matrices with determinant zero

Let $\mathbf{A}$ be a 2×2 matrix with determinant zero.

- The eigenvalues of $\mathbf{A}$ are 0 and $\operatorname{tr}\mathbf{A}$. (If $\operatorname{tr}\mathbf{A} = 0$, then 0 is a repeated eigenvalue of $\mathbf{A}$.)
- Each non-zero column of $\mathbf{A}$ is an eigenvector of $\mathbf{A}$ corresponding to $\operatorname{tr}\mathbf{A}$. To find an eigenvector corresponding to 0, use the usual method.

Generalised scalings

A **generalised scaling** of the plane is a linear transformation represented by a 2×2 matrix that has two real eigenvectors that are not scalar multiples of one another. It scales distances in the direction of an eigenvector by a factor equal to the magnitude of the corresponding eigenvalue.

Diagonalising matrices

A **diagonal matrix** is a square matrix such that each element not on its leading diagonal is zero.

A 2×2 matrix $\mathbf{A}$ is **diagonalisable** if there is a 2×2 diagonal matrix $\mathbf{D}$ and a 2×2 invertible matrix $\mathbf{P}$ such that

$$\mathbf{A} = \mathbf{PDP}^{-1}.$$

The process of finding matrices $\mathbf{D}$ and $\mathbf{P}$ that satisfy this equation is called **diagonalising** the matrix $\mathbf{A}$.

To diagonalise a 2×2 matrix $\mathbf{A}$ with two distinct real eigenvalues

1. Find the eigenvalues λ_1 and λ_2 of $\mathbf{A}$, and corresponding eigenvectors $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$, respectively.
2. Form the 2×2 matrix $\mathbf{P} = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$.
3. Find the inverse $\mathbf{P}^{-1}$ of $\mathbf{P}$.
4. Write down the equation $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

Trace formula for a 2×2 matrix

If $\mathbf{B}$ is a 2×2 matrix, and $\mathbf{P}$ is an invertible 2×2 matrix, then

$$\operatorname{tr} \mathbf{B} = \operatorname{tr}(\mathbf{PBP}^{-1}).$$

Powers of a 2×2 diagonal matrix

Let $\mathbf{D} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$. Then $\mathbf{D}^n = \begin{pmatrix} a^n & 0 \\ 0 & d^n \end{pmatrix}$ for $n = 1, 2, 3, \ldots$.

To find $\mathbf{A}^n$, where $\mathbf{A}$ is a 2×2 diagonalisable matrix and n is a positive integer

1. Diagonalise $\mathbf{A}$; that is, find a diagonal matrix $\mathbf{D} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$ and a 2×2 invertible matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$.
2. Calculate $\mathbf{D}^n = \begin{pmatrix} a^n & 0 \\ 0 & d^n \end{pmatrix}$.
3. Apply the formula $\mathbf{A}^n = \mathbf{PD}^n\mathbf{P}^{-1}$.

Useful check: The trace of $\mathbf{A}^n$ is equal to the trace of $\mathbf{D}^n$.

Systems of differential equations

A **system of differential equations** is a set of simultaneous equations, each involving the same set of unknown functions and their derivatives. Unit 11 deals only with systems of the form

$$\dot{x} = ax + by,$$
$$\dot{y} = cx + dy,$$

where x and y are functions of a variable t, and a, b, c and d are real numbers.

A system of this form is a **decoupled system** if $b = c = 0$. Otherwise, it is a **coupled system**.

In either case, the equations can be rewritten as a single matrix equation:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \qquad \text{where} \qquad \mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \ \text{and} \ \dot{\mathbf{x}} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}.$$

In this form the system is called a **matrix differential equation**.

A **solution** of the system above consists of an equation for x in terms of t and an equation for y in terms of t that simultaneously satisfy the system. The two equations can be written together as a single vector equation.

The **general solution** of the system is a solution containing arbitrary constants, such that any solution of the system is obtained by choosing particular values for the arbitrary constants.

Solving a system of differential equations means finding its general solution.

To find the general solution of a system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ has two distinct real eigenvalues

Consider the system of differential equations

$$\dot{x} = ax + by,$$
$$\dot{y} = cx + dy,$$

where a, b, c and d are real numbers such that the matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has two distinct real eigenvalues. To solve the system, proceed as follows.

1. Write down the matrix form of the system of differential equations.
2. Find the eigenvalues λ_1 and λ_2 of the matrix, and find corresponding eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$.
3. Write down the general solution, which is

   $$\mathbf{x} = Ce^{\lambda_1 t}\mathbf{x}_1 + De^{\lambda_2 t}\mathbf{x}_2,$$

   where C and D are arbitrary constants.
4. State the solution as an equation for x and an equation for y.

MST125 Unit 12 Combinatorics

Counting problems

A **counting problem** is a problem that involves finding the number of objects of a particular type or the number of ways of doing something, where the number of objects or ways is a non-negative integer.

Sets

For basic definitions relating to sets, see page 28.

A **finite set** is one that contains n elements for some non-negative integer n.

An **infinite set** is a set that is not finite.

The **size** of a finite set S is the number of elements in S. It is denoted by $|S|$.

The **complement** of a set S **in** a set U (where S is a subset of U) is the set of all elements of U that are not elements of S. It is denoted by $\overline{S}$.

Two sets are **disjoint** if their intersection is the empty set.

Two or more sets are **pairwise disjoint** if the intersection of any two of them is the empty set.

The number of subsets of an n-element set

A set containing n elements has 2^n subsets, for any non-negative integer n.

Principles of counting

The multiplication principle

If you have k successive choices to make, and the ith choice involves choosing from n_i options (for each $i = 1, 2, \dots, k$), then the total number of ways to make all k choices is

$$n_1 \times n_2 \times \cdots \times n_k.$$

The addition principle

If the finite set S can be divided into a collection of k pairwise disjoint subsets $S_1, S_2, \dots, S_k$ whose union is S, then

$$|S| = |S_1| + |S_2| + \cdots + |S_k|.$$

The subtraction principle

If the set S is a subset of the finite set U, then

$$|S| = |U| - |\overline{S}|.$$

Choosing a strategy for solving a counting problem

The following steps may help.

1. Express the problem in terms of counting elements of a set S.
2. Pick some element from S, and try to identify a sequence of simple decisions that together are equivalent to choosing this element. It may help to sketch part of a tree diagram for the problem.
3. Look at the sequence of decisions that you identified in step 2. Ask yourself what other options could be selected at each decision, and what other elements of S can be chosen in this way.
4. Repeat steps 2 and 3 until you have a good understanding of how all the elements of S can be chosen. Each time you pick a new element, try to construct a sequence of decisions that is similar to the sequences of decisions that you have already found.
5. Using your sequence(s) of decisions, try to find ways to use the counting principles to count the total number of elements of S.
   - Use the multiplication principle when you have a sequence of decisions where the number of options for each decision is not affected by decisions already made.
   - It may help to divide the problem into a few subproblems, by splitting S into a number of pairwise disjoint subsets. You may then be able to use the multiplication principle to count the number of elements in each subset, and use the addition principle to find the total number of elements in S.
   - In particular, when you have a decision that affects the number of options for subsequent decisions, you may find that you can use the addition principle together with the multiplication principle in the way described above.
   - Some decisions may involve choosing a sequence, permutation or combination. If so, then you can use the formulas on the next page for the numbers of these.
6. If it seems hard to count the number of elements of S, then think about whether S is a subset of a larger set U, where both U and the complement $\overline{S}$ of S in U are easier to count than S. If so, use the subtraction principle. In particular, the subtraction principle is often useful for problems that contain phrases such as 'at least one', or 'one or more'.

Sequences, permutations and combinations

The four different ways of choosing k objects from n objects are listed below. For permutations and combinations, $k \leq n$.

Way of choosing k objects from n objects	Meaning	Number
Sequence	Order important, repetitions allowed	n^k
Permutation	Order important, repetitions not allowed	${}^nP_k = n(n-1)\cdots(n-k+1) = \dfrac{n!}{(n-k)!}$.
Combination	Order not important, repetitions not allowed	${}^nC_k = \dfrac{n(n-1)\cdots(n-k+1)}{k!} = \dfrac{n!}{(n-k)!\,k!}$
Multiset	Order not important, repetitions allowed	(not covered in MST125)

Arrangements and selections

An **arrangement** is a choice of k objects from n objects in which order is important (that is, it is a sequence or a permutation).

A **selection** is a choice of k objects from n objects in which order is not important (that is, it is a combination or a multiset).

Probability

An **outcome** is a possible result of an experiment or process.

The **sample space** of an experiment or process is the set of all possible outcomes.

An **event** is a subset of the sample space.

The probability of an event

For a sample space S in which each outcome is equally likely, the probability $P(E)$ that the event E happens is given by

$$P(E) = \frac{|E|}{|S|}.$$

Recurrence relations and closed forms

A **recurrence relation** for a sequence is an equation that expresses each term a_n from some term onwards in terms of one or more previous terms. We say that the sequence **satisfies** the recurrence relation.

A **recurrence system** for a sequence consists of the following:

- a recurrence relation for the sequence
- the values of enough initial terms to enable the remaining terms to be found by using the recurrence relation
- the range of values of the index variable for which the recurrence relation applies.

A **closed form** for a sequence is a formula that defines the general term a_n as an expression involving the index variable n.

Solving a recurrence system means finding a closed form for the sequence that it specifies.

A **general solution** of a recurrence relation is a closed form containing one or more arbitrary constants, such that any sequence satisfying the recurrence relation has a closed form obtained by choosing particular values for the arbitrary constants.

First-order recurrence systems

A **first-order recurrence relation** for a sequence is an equation that expresses each term x_n, other than the first, in terms of the previous term x_{n-1} and no other terms.

A **first-order recurrence system** for a sequence consists of the following:

- the value of the first term of the sequence
- a first-order recurrence relation for the sequence
- the range of values for the index variable in the recurrence relation.

A **first-order recurrence sequence** is a sequence that has a first-order recurrence relation.

A recurrence relation of the form

$$x_n = rx_{n-1} + d,$$

where r and d are constants with $r \neq 0$, is a special case of a **linear first-order constant-coefficient recurrence relation**. (It is a special case because, in general, the recurrence relation could include a function of n in place of the constant d.) The term **linear** refers to the fact that the recurrence relation involves no powers of x_{n-1} other than x_{n-1} itself. The term **constant-coefficient** refers to the fact that the coefficient of x_{n-1} (that is, r) is a constant, rather than a function of n.

In MST125 we refer to a recurrence relation of the form above simply as a **first-order recurrence relation**, and similarly for recurrence systems and recurrence sequences.

The **parameters** of a first-order recurrence sequence are its first term a and the constants r and d in its recurrence relation $x_n = rx_{n-1} + d$.

To find a closed form for a first-order recurrence sequence

Suppose that the sequence $x_1, x_2, x_3, \dots$ is given by the recurrence system

$$x_1 = a, \quad x_n = rx_{n-1} + d \quad (n = 2, 3, 4, \dots),$$

where $r \neq 1$. To find a closed form for this sequence, proceed as follows.

1. Find the second term x_2 of the sequence.
2. Write down the general solution of the recurrence relation, which is

   $$x_n = Cr^{n-1} + D \quad (n = 1, 2, 3, \dots),$$

   where C and D are constants.
3. Find the values of C and D by using the values of the initial terms x_1 and x_2 to obtain two simultaneous equations in C and D and solving them.

To find a first-order recurrence system for a sequence of counting problems

Suppose that the sequence (x_n) counts the number of objects with some property that depends on n. To try to find a first-order recurrence system for (x_n), proceed as follows.

1. Find the first two or three values of x_n by listing the objects. This will give you a feeling for the problem.
2. It may also help to try to find a way of visualising the problem using a diagram.
3. Try to express the number of objects counted by x_n in terms of the number of objects counted by x_{n-1}. That is, try to find a recurrence relation satisfied by the sequence (x_n).
   - Make sure that you give a clear, logical argument (a proof) that the recurrence relation is correct.
   - It may help to divide the problem into several cases and use the addition principle.
   - Sometimes it may help to count the objects not being counted by x_n, and use the subtraction principle.
4. As a check, determine whether the recurrence relation works for the first few values of n.
5. Specify the required recurrence system by giving the following:
   - the value of the first term in the sequence
   - the recurrence relation
   - the range of values for the index variable n.

Second-order recurrence systems

A **second-order recurrence relation** for a sequence is an equation that expresses each term u_n, other than the first two, in terms of the previous two terms u_{n-1} and u_{n-2} and no other terms.

A **second-order recurrence system** for a sequence consists of the following:

- the values of the first two terms of the sequence
- a second-order recurrence relation for the sequence
- the range of values for the index variable in the recurrence relation.

A **second-order recurrence sequence** is a sequence that has a second-order recurrence relation.

A **linear second-order constant-coefficient homogeneous recurrence relation** is a recurrence relation of the form

$$u_n = pu_{n-1} + qu_{n-2},$$

where p and q are constants, with $q \neq 0$.

The term **linear** refers to the fact that the recurrence relation involves no powers of u_{n-1} or u_{n-2} other than u_{n-1} and u_{n-2} themselves.
The term **constant-coefficient** refers to the fact that the coefficients of u_{n-1} and u_{n-2} (that is, p and q) are constants, not functions of n.
The term **homogeneous** refers to the fact that all the terms on the right-hand side of the recurrence relation involve u_{n-1} or u_{n-2}.

In MST125 we refer to a recurrence relation of the form above simply as a **second-order recurrence relation**, and similarly for recurrence systems and recurrence sequences.

To find a closed form for a second-order recurrence sequence

Suppose that the sequence $u_0, u_1, u_2, \dots$ is given by the recurrence system

$$u_0 = a, \quad u_1 = b, \quad u_n = pu_{n-1} + qu_{n-2} \quad (n = 2, 3, 4, \dots).$$

To find a closed form for this sequence, proceed as follows.

1. Write down and solve the **auxiliary equation** for the recurrence relation, which is

   $$r^2 - pr - q = 0.$$

   There may be distinct real solutions, say $r = \alpha$ and $r = \beta$, or a single (repeated) real solution, say $r = \alpha$, or no real solutions.

2. Write down the general solution of the recurrence relation, using the table below, where A and B are constants.

Solutions of auxiliary equation	General solution
Distinct real solutions α, β	$u_n = A\alpha^n + B\beta^n$
Single real solution α	$u_n = (A + Bn)\alpha^n$
No real solutions	(not covered in MST125)

3. Find the values of A and B by using values of the initial terms u_0 and u_1 to obtain two simultaneous equations in A and B and solving them.

The Fibonacci sequence

The **Fibonacci sequence** is the sequence (F_n) given by the second-order recurrence system

$$F_1 = 1, \quad F_2 = 1, \quad F_n = F_{n-1} + F_{n-2} \quad (n = 3, 4, 5, \dots).$$

Sometimes it is convenient to include a 0th term, $F_0 = 0$.

Binet's formula

A closed form for the Fibonacci sequence (F_n) is

$$F_n = \frac{1}{\sqrt{5}}(\phi^n - \psi^n) \quad (n = 0, 1, 2, \dots),$$

where $\phi = \frac{1}{2}(1 + \sqrt{5})$ (the **golden ratio**) and $\psi = \frac{1}{2}(1 - \sqrt{5})$.

Binet's approximation

For each $n = 0, 1, 2, \dots$, the nth term F_n of the Fibonacci sequence is the nearest integer to $\dfrac{\phi^n}{\sqrt{5}}$, where $\phi = \frac{1}{2}(1 + \sqrt{5})$.

To find a second-order recurrence system for a sequence of counting problems

Try using a strategy similar to the one for a first-order recurrence system on page 120.

Index

BOOK A

Unit 1 Key techniques

Unit 2 Mathematical typesetting
(Provided as a separate item)

Unit 3 Number theory

BOOK B

Unit 4 Conics

Unit 5 Statics

Unit 6 Geometric transformations

BOOK C

Unit 7 Topics in calculus

Unit 8 Differential equations

Unit 9 Mathematical language and proof

BOOK D

Unit 10 Dynamics

Unit 11 Eigenvalues and eigenvectors

Unit 12 Combinatorics

SUP 04700 7

Cover image designed by Andrew Whitehead